UNDERSTANDING YOUNGER AND OLDER ADULTS' PERCEPTIONS OF HUMANOID ROBOTS: EFFECTS OF FACIAL APPEARANCE AND TASK

A Thesis
Presented to
The Academic Faculty

by

Akanksha Prakash

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Psychology

Georgia Institute of Technology
August 2013

UNDERSTANDING YOUNGER AND OLDER ADULTS' PERCEPTIONS OF HUMANOID ROBOTS: EFFECTS OF FACIAL APPEARANCE AND TASK

Approved by:

Dr. Wendy A. Rogers, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Arthur D. Fisk
School of Psychology
Georgia Institute of Technology

Dr. Leslie A. DeChurch
School of Psychology
Georgia Institute of Technology

Date Approved: April 17, 2013

To my late grandparents,
Sri Raghunath Prasand and Smt. Rajdevi

ACKNOWLEDGEMENTS

This research was supported in part by a grant from the National Institute of Health (National Institute of Aging) Grant P01 AG17211 under the auspices of the Center for Research and Education on Aging and Technology Enhancement (CREATE). I would like to thank my advisor, Dr. Wendy A. Rogers, for her mentoring and support at every stage of this study. I am also grateful to Dr. Arthur D. Fisk and Dr. Leslie A. DeChurch for their guidance and willingness to serve on my committee. I wish to extend special thanks to all the members of the Human Factors and Aging Laboratory for their help and cooperation. Finally, I would like to thank my family and friends for their unconditional love and understanding.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

In human-human interaction, facial appearances influence formation of initial impressions (Bar, Neta, & Linz 2006; Masip, Garrido, & Herrero, 2004; Zebrowitz & Monteparo, 2008). However, a tendency to over-interpret facial cues is also not uncommon (Zebrowitz & Rhodes, 2004). Moreover, initial impressions, even when inaccurate, influence the behavior of the perceiver (e.g., Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; Riggio, Widaman, Tucker, & Salinas, 2010). It is an open question if similar socio-cognitive processes underlie human perceptions and behaviors in human-robot interaction when the robot's form and functionality resemble a human's.

Currently, many humanoid robots are being designed with the assumption that the humanness of the robot would ease and enhance the nature of human-robot interaction because humans are hard-wired for human-human interactions (Blow, Dautenhahn, Appleby, Nehaniv, & Lee, 2006). Despite the focus on the creation of somewhat to highly human-looking robots to provide assistance with various tasks, there remain gaps in our understanding of the perceptions that humanoid faces evoke in the user.

Understanding user perceptions would help design robots that are better suited for the target user group. Thus, one of the primary goals of this study was to investigate how initial perceptions of robots are influenced by the extent of humanness of the robot's face, particularly when the robot is intended to provide assistance with tasks in the home that are traditionally carried out by humans. Moreover, although robots have the potential to help both younger and older adults, there is limited knowledge on how the two age

groups' perceptions of robot humanness compare with each other. Therefore, an additional goal was to examine if younger and older adults differed in their perceptions.

At a general level, a mixed human-robot facial appearance was evaluated less positively than a highly human-looking or a highly robot-looking appearance. This trend was observed in ratings on the measures of perceived usefulness, likeability, and trust for both younger and older adults. This finding seems aligned with the uncanny valley theory (Mori, 1970), implying that a robot face that partially imitates a human appearance evokes less positive perceptions than a more mechanical or a completely human-like robot face. However, one of the caveats of the earlier research on uncanny valley theory was the ill-defined context in which robot appearances were evaluated (e.g., MacDorman, 2006; MacDorman & Ishiguro 2006). This caveat was addressed in the current study by asking participants to imagine interacting with the robot in specific task contexts.

When the task was taken into account, the trends in perceptions were more complex and deviations from the uncanny valley pattern were observed. For example, robot (mechanical) appearance was evaluated more positively than the mixed appearance for chores, social, and personal care tasks. However, for decision-making task, mean ratings for robot appearance were comparable to those for the mixed appearance.

Prior research on robot appearance that did take robot task into consideration did not assess the underlying reasons for the preference of one appearance over the other (e.g., Goetz et al., 2003). The multi-method approach used in the current study identified not only the patterns of perceptions across different appearances but also the reasons that influence the formation of such perceptions. The interview data revealed that participants varied their evaluation criteria for robot appearance across different tasks based on the attributes of the task. For the decision-making task, the appearance that evoked perceptions of intelligence, smartness, or wisdom was preferred for assistance. Perceptions of “cuteness” or “friendliness”, which were frequently mentioned as reasons for a general preference of the mechanical appearance were not held important when evaluating assistance for a cognitively demanding task such as decision-making.

Age-related differences in preferences of robot humanness were also observed. Older adults showed a higher inclination toward human-looking appearance of robots whereas younger adults’ preferences were more distributed across the levels of humanness. An appearance with mixed human-robot features was more likely to be rejected by older adults than by younger adults, and the difference was most striking for a decision-making task. Besides the humanness of the robot face, perceptions of robot appearances were also influenced by factors such as robot gender, specific facial features/aesthetics, expressiveness, perceived personality, and perceived capability.

Overall, the results of this research clearly indicated that people’s perceptions of robot faces vary as a function of robot humanness. Additionally, the nature of task leads to intra-individual differences whereas age-cohort acts as a source of inter-individual differences in the perceptions of humanoid robots.

CHAPTER 1

INTRODUCTION

Why Facial Appearance Matters

Faces are important for social interaction by serving not only as a marker of identity to distinguish visually one person from another, but also as a canvas for the display of non-verbal social cues. From an ecological perspective, such cues have been found to fulfill a two-fold purpose (Zebrowitz & Monteparo, 2008). Firstly, facial cues play an adaptive role in human-human interaction by communicating to the observer the condition, intention, or need of the observed. For example, a baby's cute face signals its vulnerability and evokes protective instincts in the care-taker (Zebrowitz, 1997). Secondly, meanings of such cues are learned and generalized to decipher new faces in unfamiliar contexts (e.g., in forming impressions about the personality of a stranger with a baby-like face; Masip, Garrido, & Herrero, 2004).

People's initial impressions of others are often influenced by perceptions of facial appearance (Bar, Neta, & Linz 2006; Masip, et al., 2004; Zebrowitz & Monteparo, 2008). However, people tend to "over-read" or over-generalize the cues emanating from a face. "Attractiveness halo" is one of the most common phenomena resulting from such over-generalizations (Zebrowitz & Monteparo, 2008). People with attractive faces are rated more positively on a wide variety of dimensions; for example, they are perceived as more intelligent, healthy, and sociable than people with less attractive faces (Zebrowitz & Rhodes, 2004). Baby-faced overgeneralization is another well-documented instance whereby people ascribe certain personality traits that are characteristic of a child to individuals whose facial structures resemble that of a baby (Berry & Landry, 1997; Masip et al., 2004; Zebrowitz & Monteparo, 2008).

Initial impressions about a person, even when inaccurate, influence the behavior people adopt toward the person. For instance, people are more likely to date a person whom they perceive to be physically more attractive (Riggio, Widaman, Tucker, & Salinas, 2010; Walster, Aronson, & Abrahams, 1966). In an altogether different context, it has been found that if a criminal case involves a white victim, the more the defendant's face is perceived to have stereotypical characteristics of a black person, the more likely he is to be sentenced to death (Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006). Furthermore, facial appearances also affect perceptions of competence of political candidates, thereby influencing voting choices (Todorov, Mandisodza, Goren, & Hall, 2005).

Given the research on the perception of human faces, it is possible that similar socio-cognitive processes might be at play in the formation of initial impressions about human-looking robots, particularly in situations where human-robot interaction resembles human-human interaction. Although a robot is inherently a machine, when given a human-like appearance, and designed to perform tasks originally carried out by humans, people may *over*-generalize their understanding of human facial cues to build expectations about its behavior and capabilities, which may further impact their own behavior toward and acceptance of the robot in question.

Prevalence of Humanoid Robots

Traditionally, the use of robots is identified with military, manufacturing, and space-research domains. However, robot applications are now extending to domestic, healthcare, and entertainment settings. Thus, although various service robots are being designed for personal use in the home environment, the target users, in general, are expected to be naïve to the engineering complexity of the robotic system. Many robotics researchers believe that because humans are hard-wired or more accustomed to human-human interaction, giving robots human form and functionality will enhance and ease the quality of human-robot interaction (Blow, Dautenhahn, Appleby, Nehaniv, & Lee, 2006). The increasing prevalence of "humanoid" robots is, therefore, not surprising. Humanoid is an umbrella term for robots that have some human resemblance. Extreme resemblance to a human marks the category of "android robots". Thus, androids are humanoid robots that are designed to be almost indistinguishable from human beings in appearance and behavior (MacDorman & Ishiguro, 2006).

Despite the emphasis on the design of humanoid robots, the existing humanoids vary to a great extent and not in a systematic fashion. Almost all humanoid robots have heads, but there are variations in the head shapes (e.g., round versus rectangular, wide versus narrow) and also in the composition of facial features. Therefore, researchers are trying to develop criteria that can enable the assessment of a robot's human-likeness. A study on the design of humanoid heads found that 62% of the variation in the perception of human-likeness of humanoid heads was accounted for by the presence of particular facial features (Disalvo, Gemperie, Forlizzi, & Kiesler, 2002). Nose, eyelids, and mouth were found to be the facial features that provide most enhancements to a robot's human-likeness. The study also found that if the head was wider than it is tall, the robot was perceived as more robot-like (and therefore, less human-like) in appearance. Moreover, the appearance became less human-like if the proportion of head-space for forehead, hair, or chin was reduced.

Such criteria can be used to evaluate the human-likeness of a robot's face. However, the intriguing question, which remains only partially answered so far, is how people perceive different robots that vary in their extent of humanness. Although attempts have been made to understand the effects of humanness, the trends revealed in the existing literature on the perception of robots do not converge to give a clear picture. Thus, one of the primary goals of this research is to understand people's perceptions of robot faces across a range of humanness in appearance.

The Uncanny Valley Theory

Studies on the perception of humanoid robots are often aimed at investigating the validity of the uncanny valley theory (Mori, 1970; translated by MacDorman & Minato, 2005). The uncanny valley theory is a popular theory that tries to relate human-likeness of a robot with the level of familiarity evoked in the human observer. This relation is posited to be curvilinear. According to the theory, as a robot appears more and more human-like, people's familiarity with it increases until a point where this relationship ceases. Beyond this critical point, the appearance of the robot increases in human-likeness but the appearance no longer evokes a feeling of familiarity. The robot instead is perceived as strange or eerie. If a robot's human-likeness is further increased to almost entirely match the appearance of a human, familiarity will rise again and will be maximized when the robot cannot be distinguished from a healthy person. The region of dip in familiarity with increasing human-likeness is referred as the uncanny valley (see Figure 1.1).

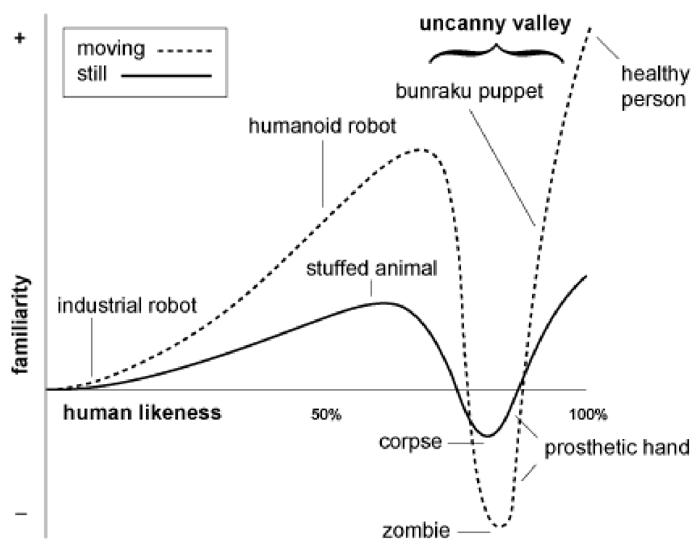


Figure 1.1. Mori's hypothesized uncanny valley diagram (translated by MacDorman & Minato, 2005)

The main limitation of the uncanny valley theory is that it was formulated based on anecdotal examples. It has been claimed that at its conception, the theory was not experimentally verified. Instead, it was proposed by Mori as a generalization of his experiences with prosthetic hands, mannequins, and robots (MacDorman & Ishiguru, 2006). However, despite its non-empirical basis, it has triggered plentitude of research, probably due to its historical significance as one of the earliest theories on human perception of human-like robots (e.g., Chen, Russel, Nakayama, & Livingstone, 2010; Geller, 2008; Groom et al., 2009; MacDorman & Ishiguru, 2006; MacDorman, Green, Ho, & Koch, 2009; Schneider, Wang, & Yang, 2007; Thompson, Trafton, & McKnight, 2011).

Empirical Investigations of the Uncanny Valley Theory

To assess if the uncanny valley can be plotted as Mori hypothesized, MacDorman & Ishiguro (2006) morphed robot faces onto human faces in different proportions and used them as stimuli to gauge people's perceptions. The study was conducted with Indonesian participants through computer-based questionnaires. Participants saw the still images of the generated faces and rated them on scales for human-likeness (1=very mechanical; 9 = very humanlike), familiarity (1=strange; 9= very familiar), and eeriness (0=not eerie; 10= extremely eerie). People's familiarity ratings when plotted against perceived robot human-likeness ratings resulted in the uncanny valley pattern.

In a later study, a different set of Indonesian participants viewed *video clips* of 13 existing robots and a human, and rated each robot and the human on human-likeness, familiarity, and eeriness (MacDorman, 2006). In this case, the plots of human-likeness versus familiarity did not reveal a U-shaped plot predicted by uncanny valley theory. Due to such discordant findings with the varying nature of stimuli, it remains unclear as to what factors are most impactful in the initial perception of humanoid robot faces.

Although still images have limitations of not displaying subtle, dynamic movements and expressions that may further impact perceptions of the beholder, videos are also not without constraints. The apparent caveat in the MacDorman (2006) study was the use of a wide range of robots performing dissimilar actions in different settings. It is, therefore, not possible to decipher how participants' appraisal of the robots' activities and environments informed their impressions of the robots' appearances. Moreover, some robots had voices whereas others were without which was an added confound in the design of the study. Thus, due to lack of systematic manipulation of relevant variables, participants' ratings of the videos did not elucidate what characteristics of the robot features and actions were being attended to and were influencing the formation of perceptions.

In both of these studies, the participants merely viewed the robots and rated them without any (actual or imagined) context of interaction with the robots. It is plausible that if the viewers were facilitated to perceive the robots as performing a relevant task for them, their ratings of familiarity would be more indicative of their attitudes. Additionally, familiarity in itself is an ambiguous construct and is not informative of people's preferences. High familiarity may not necessarily imply liking or acceptance. Similarly, low familiarity may not always imply disliking or rejection. Probably, because the translation of Mori's original article (Mori, 1970; translated by MacDorman & Minato, 2005) used the term "familiarity", researchers investigating the uncanny valley continued to measure how familiar people find robots based on their appearances. However, more specific measures need to be used to gain clearer insights into people's perceptions of humanoid robots.

Age-Related Differences in the Perception of Robots

The plethora of research on the uncanny valley theory notwithstanding, there is lack of concrete evidence to infer the effects of a robot's human-likeness on the human perceiver. Some major methodological limitations of the research on uncanny valley were discussed in the previous section. Another primary limitation is the narrow range of potential users being focused upon. Most research on robot appearance has involved only young adult participants. To test the validity of any general theory on robot perception, a broader age-range of participants should be considered.

Considering different age groups is essential not only to enhance a theoretical understanding of robot perception but also to improve the design of robots so they are more acceptable to the target users. From the perspective of application, robots have the potential to support people's independence and well-being. They can specifically assist older adults with various home-based tasks, so they can continue living independently in their homes (Beer et al., 2012). The limited human-robot interaction (HRI) studies conducted with older adults have shown that although older adults have less experience with robots (Prakash et al., 2013), they have expectations and opinions about robot appearance, particularly in terms of size. (Ezer, 2008; Wu, Fassert, & Rigaud, 2012). However, we are limited in our knowledge of how older adults perceive highly human-looking robots (androids) in comparison to less human-looking ones (humanoids).

There are also gaps in our understanding of younger adults' perceptions of humanoid robots in comparison to androids. Researchers often focus on comparing perceptions toward mechanical appearance (devoid of any human features) with humanoid robots. For example, a study conducted with university undergraduates concluded that younger adults in general showed preference for human-like appearance of robots although large individual differences in preferences were noted (Walters, Syrdal, Dautenhahn, Boekhorst, & Koay, 2008). However, even the most human-like appearance manipulated in the study had some human features but was not close to a human appearance. It remains unclear if younger adults' perceptions would change and in which direction if the appearance were made more human-looking up to the point that it matches the appearance of a human.

Although more research has been conducted with younger adults than with older people, we have limited knowledge of both the age groups' perceptions of humanoid to android appearances for robots. Understanding perceptions of a broad range of users can guide the design of robots that are acceptable to the target user. Moreover, robots are assistive technologies designed to perform some or many tasks. Thus, people's perceptions of robots and robot appearances need to be assessed in the context of a task.

Perceptions in the Context of Robot Task

The home setting is increasingly being considered as a large market for service robot applications. A wide range of robots of varied appearances are currently under development that may potentially assist with everyday living tasks (for a review, see Smarr, Fausset, & Rogers, 2011). Such assistive robots have the potential to support people's independence and well-being, and can be specifically beneficial in helping people age successfully in their own homes.

There are many tasks that people must perform to maintain their independence and health, including self-maintenance, instrumental, and enhanced activities of daily living (Lawton, 1990; Rogers, Meyer, Walker, & Fisk, 1998). Self-maintenance activities of daily living (ADLs) include the ability to toilet, feed, dress, groom, bathe, and ambulate. Instrumental activities of daily living (IADLs) include the ability to successfully use the telephone, shop, prepare food, do the housekeeping and laundry, manage medications and finances, and use transportation. Enhanced activities of daily living (EADLs) include participation in social and enriching activities, such as learning new skills and engaging in hobbies. Age-related changes in physical, perceptual, and cognitive abilities may make performing these tasks more difficult or challenging for older adults.

Even for the younger people, robots may play a beneficial role by saving their time and effort. Moreover, high workload in people's professional lives may prevent them from regularly taking care of some or many household activities. Well-designed, functional robots can facilitate timely maintenance of the home and its surroundings, provide entertainment and companionship, help in solving cognitively challenging, intellectual problems, and also assist in personal care tasks, if needed.

There is early evidence suggesting that people's attitudes toward robot assistance vary with the task (Broadbent et al., 2011; Prakash et al., 2013, Smarr et al., 2012). Older adults are selective in their preferences of robot assistance over human assistance (Smarr et al., 2012). The selectivity is determined by the nature of the home-based task. In general, there is higher attitudinal acceptance of robots for assistance with IADLs (e.g., chores), followed by EADLs (e.g., learning a new skill). Older adults are least open to robot assistance for ADLs (e.g., bathing). However, these findings may be influenced by the specific type of robot being considered.

Thus, although task seems to determine the level of acceptance of a robot, less is known about the effect robot task will have on how people evaluate the extent of humanness in robot appearance. Are there tasks for which a highly human-looking robot would be evaluated more positively than less human-looking appearances? Are there tasks for which the trend would reverse? In sum, the open question is how task and human-likeness jointly impact perceptions of robots?

It has been suggested that an appropriate match between a robot's appearance and its task can improve people's acceptance of the robot (Goetz, Kiesler & Powers, 2003). People are likely to prefer human-looking robots to perform jobs that entail more social skills (e.g., sales representative, aerobics instructor) but greater preference would be shown for machine-looking robots for jobs less social in nature (e.g., customs inspector, security guard).

It is worth noting that the human-like robot faces used in the Goetz et al. (2003) study were not fine imitations of human faces. The stimuli used as human-looking faces can be considered more human-looking than the machine-looking faces employed in the study. Nonetheless, they did not resemble the appearance of a person in that the faces were simplistic, cartoon-like renditions of human facial shape, features, and hair, and less sophisticated in details. Thus, though informative about the interactive effect of robot appearance and task, the Goetz et al. (2003) study does not provide insights into people's perceptions of very human-looking robots. For instance, if a robot were designed to look indistinguishable from a human, which tasks would it be most preferred for?

Robot appearance and task also had an interaction effect when the robot played the role of a co-worker in a work environment (Hinds, Roberts & Jones, 2004). People felt more responsible when working with a machine-looking robot than when working with a human-looking robot, particularly when the robot was in a subordinate position. Based on this finding, Hinds et al. (2004) suggested that robots should be made mechanical-looking when assisting in environments where personal responsibility is important. However, the researchers used two extreme manipulations of robot appearance such that the mechanical looking robot did not have a human form whereas the human-looking robot looked like a white male. Thus, their study did not unveil the impact of intermediate human-robot appearance; that is, how would a robot with mixed human-robot features be perceived? Moreover, how would perceptions of such a robot be influenced by its task?

Measuring Perceptions

What Do "Perceptions" Comprise?

Initial perceptions of a robot can be based on different kinds of appraisals people make, such as, how useful it is perceived to be, how much trust it evokes, how likeable it seems, and how anxious it makes them feel. These appraisals seem related but the strengths of the inter-correlations are unknown. People are also likely to ascribe relative importance to every factor. For example, if a robot is considered useful even if it is not liked, which factor will have a greater impact on people's overall perception of the robot?

Some of the common measures that have been used in studies investigating the uncanny valley theory are: affect evoked such as fear and anxiety, attractiveness versus repulsiveness, familiarity, likeability, and perceived eeriness (see Appendix A for a reference list). Each of these measures informs about a particular constituent of perceptions; however they cannot independently provide a complete picture of perception-formation. The need is to evaluate perceptions on multiple dimensions together for a holistic understanding of *attitudinal* acceptance of robots. Attitudinal acceptance, defined as the users' positive evaluation or beliefs about the technology, has been argued for as a precursor to behavioral acceptance, that is, the users' actions in using the product or technology (Davis, 1989).

Variables Assessed in Technology Acceptance Models

Robots are advanced technology and the general technology acceptance models can be a starting point for the conceptualization of robot acceptance. The Technology Acceptance Model (TAM; Davis, 1989) is the most widely recognized model of technology acceptance. The TAM was developed to understand prospective expectations about information technology usage. The model proposes two main variables that affect acceptance: perceived usefulness and perceived ease of use. There is strong empirical support for the TAM (Venkatesh & Davis, 2000; Venkatesh, Morris, Davis, & Davis, 2003), in part due to its ease of application to a variety of domains. However, the model's simplicity has evoked some criticism (Bagozzi, Davis, & Warshaw, 1992) that has led to the development of other models. For example, the Unified Theory of Acceptance and Use of Technology (UTAUT) Model (Venkatesh, et al., 2003) was developed with the intent of unifying a large number of acceptance models. UTAUT posits that technology acceptance may be determined by the following constructs: performance expectancy, effort expectancy, social influence, and facilitating conditions. An alternative model, the Technology-to-Performance Chain Model (TPC; Goodhue & Thompson, 1995), asserts that technology adoption is impacted by the technology's utility and its fit with the tasks it is designed to support (referred to as task-technology fit).

Robots differ from other technologies in certain aspects, and the existing technology acceptance models cannot provide a complete picture for robot acceptance if applied without modifications. For instance, a personal robot is an *embodied* agent with social capabilities and social presence, and the expectation for it is to work in a collaborative manner with the user (Walters, 2008). This may heighten the importance of its appearance or human-likeness, the characteristics of tasks it performs, and the affect it evokes in the user. TAM and UTAUT do not include variables of appearance, task characteristics, and affect. TPC model is oriented toward information technology and, even with a task-technology fit dimension, may not be suitable as it is for an embodied agent that has a social presence outside of the computer system.

It is also worth considering that robots have been a topic of science fiction literature and film for decades. Rosie from the Jetsons, C3P0 and R2D2 from Star Wars, and Robby the Robot from Forbidden Planet, are all beloved science fiction characters and they have influenced the way in which the general public thinks about robotics. Likewise, fictional robots in antagonistic roles such as the Terminator propagate a negative image of robots. Thus, media exposure may create preconceived expectations about robots, even for individuals who have never interacted with a robot directly. In fact, people do have ideas or definitions of what a robot should be like (Ezer, 2008). Pre-existing ideas about robots may lead to evaluations of an existing robot against criteria based on one's expectations. Violations of expectations are likely to negatively impact acceptance.

Robot Attitude Scales

Some psychological scales have been developed to measure people's perceptions of robots. Of these, the most widely recognized scales are the Negative Attitude towards Robots Scale (Nomura, Kanda, Suzuki, Kato, 2004; Nomura, Suzuki, Kanda, & Kato, 2006a) and Robot Anxiety Scale (Nomura, Suzuki, Kanda, & Kato, 2006b) which are used to gauge psychological reactions evoked in humans by robots. Abbreviated as NARS and RAS respectively, use of these scales can delineate to what extent people feel unwilling to interact with a robot due to arousal of negative emotions or anxiety.

The NARS assesses negative attitudes toward robots considering three dimensions: interaction with robots, social influence of robots, and emotional interactions with robots. RAS also has three dimensions or sub-scales: anxiety toward communication capability of robots, anxiety toward behavioral capability of robots, and anxiety toward discourse with robots. It can be used to assess state-anxiety in real or imaginary interactions with robots. The limitation associated with NARS and RAS scales is that they focus only on negative affect and lack measures of positive evaluations of the robot and interactions with it. Moreover, the scales do not provide any understanding of the underlying cause of negative affect toward robots. For instance, anxiety toward a robot may result from participants' mental-models or stereotypes against robots. However, it can also be triggered due to lack of familiarity with robots in general, and thus can be eased over time as participants become more accustomed to the robots.

More recently designed scales are oriented toward both negative and positive attitudes toward robots. Robot Attitude Scale (also abbreviated as RAS; Broadbent et al., 2009) is one such scale in which a robot is rated from 1 to 8 on 11 dimensions: safe–dangerous, reliable–unreliable, friendly-unfriendly, simple–complicated, useful–useless, strong–fragile, interesting–boring, trustworthy–untrustworthy, advanced–basic, easy to use–hard to use, and helpful–unhelpful. Similarly, the Almere model, an adaptation of UTAUT, is aimed at understanding older adults' acceptance of assistive social robots (Heerink, Krose, Evers, & Wielinga, 2010) and has 9 constructs: anxiety, attitude towards technology, facilitating conditions, intention to use, perceived adaptiveness, perceived enjoyment, perceived ease of use, perceived sociability, perceived usefulness, social influence, social presence, trust, and use. These scales and models are useful developments in the space of human-robot interaction. However, their purpose is limited to identifying general trends without conclusive explanations for *why* people hold certain perceptions.

Overview of Present Research

The current literature on perception of humanoid robots has identified important variables such as robot's appearance, task, and user characteristics that can affect perceptions of robots. However, in most cases, these variables are defined and manipulated differently and are often studied in isolation from other variables. Thus, the information gauged from such studies is difficult to integrate into a holistic understanding of people's initial perceptions of robots. This study was designed to address this gap in the existing research. The specific questions explored in the present study are as follows:

1. How do people's perceptions of robot faces vary for a range of humanness in appearance?
2. Do perceptions of robots of different levels of humanness vary across tasks?
3. Do younger and older adults differ in their perceptions of robots?

In the present study, four tasks were selected and individuals were instructed to imagine the robot assisting them with the completion of each task. These tasks exemplified each of the three categories of daily living activities (i.e., ADL, IADL, and EADL). The ADL and EADL categories were represented by one task each. The IADL category was instantiated through two examples to represent the level of cognitive demand. Activities such as chores in the home are IADLs with low level of cognitive demand; finance management and medication management are tasks that can impose high cognitive load on the individual.

Four dependent variables were used to assess people's perceptions of robots: likeability, anxiety, trust, and perceived usefulness, in the four task contexts. These variables were selected to represent the range of variables assessed in the literature, and capture both affective (i.e., likeability, anxiety) and cognitive (i.e., perceived usefulness) components of individuals' attitudes (affective events theory; Weiss & Cropanzano, 1996). Trust incorporates both affect and cognitive components (Lewis & Weigert, 1985; McAllister, 1995). Likeability can indicate whether people would generally like a robot that has a certain level of humanness to assist them with a particular task. Trust in a robot is a predictor of its acceptance and is a dimension used in the most recent robot attitude scales (Broadbent et al., 2009, Heerink, et al., 2010). Perceived usefulness was measured as it is one of the main variables in the technology acceptance model (Davis, 1989). Anxiety is frequently used in the assessment of human robot interaction. For example, the Robot Anxiety Scale (RAS; Nomura et al., 2006b) is solely focused on the measurement of anxiety. The more recently developed Almere model also includes a measure of anxiety (Heerink et al., 2010). Thus, the goal behind using multiple dependent variables was to gain a holistic understanding of people's perceptions of robots across a range of dependent variables.

In the current study, participants imagined being assisted by robots in the aforementioned task contexts. For every task, they rated different robot pictures shown on a computer screen. At the end of this rating task, participants were briefly interviewed about their preferences for one robot face over the others. Additionally participants filled out various questionnaires at different points in the study. The goal behind using a combination of these methods was to assess the trend in people's reactions to robot faces in different task contexts and to understand the underlying reasons for the same.

CHAPTER 2

METHOD

Participants

The participants were 32 younger adults (18 females) and 32 older adults (19 females). The younger adults ranged in age from 18 to 23 (M = 20.16, SD = 1.42); the older adults were between the ages of 65 and 75 (M = 70.09, SD = 3.07). All younger adult participants were recruited from the Georgia Institute of Technology undergraduate population, and received credit for participation as a course requirement. The older adults were recruited from the Human Factors and Aging Laboratory database and were compensated $36 for their participation. A majority (78%) of older adults reported having a college or higher degree.

The younger and older adult samples were diverse in race/ethnicity. In the younger adult group, 62.5% reported themselves as White/Caucasian, 6.25% as Black/African American, and 28.12% as Asian. One younger adult did not report her race. In the older adult group, 62.5% reported themselves as White/Caucasian, 34.37% as Black/African American, and 3.12% as multi-racial.

Participants also provided general information about their health. They responded to the questionnaire item “In general, would you say your health is…” on a 5-point scale (1= poor; 3 = good, 5 = excellent). On average, both younger adults (M = 3.81; SD = 0.86) and older adults (M = 3.66; SD = 0.94) reported having good health.

All participants completed three ability tests during the study: Reverse Digit Span test for memory span (Wechsler, 1997), Digit Symbol Substitution test for perceptual speed (Wechsler, 1997), and Shipley vocabulary test for verbal ability (Shipley, 1986; Table 2.1). Independent samples t-tests were conducted to analyze the differences between the means of younger and older adults' abilities. The ability tests descriptives and t-statistics are reported in Table 2.1.

Significant age differences were found for the digit symbol substitution and Shipley vocabulary tests ($p < 0.05$). Specifically, compared to older adults, younger adults had higher perceptual speed but lower verbal ability. These results were consistent with the cognitive aging research (e.g., Czaja et al., 2006). Every participant's score on each test was within 3 standard deviations from the mean of their group (younger or older adults).

Table 2.1

Participants' Ability Test Scores

	Younger Adults		Older Adults		
	M	SD	M	SD	t value
Digit-Symbol Substitution[a]	76.87	13.41	52.78	13.68	7.11*
Reverse Digit Span[b]	7.44	1.97	6.94	2.27	0.94
Shipley Vocabulary[c]	30.22	3.73	33.56	4.25	-3.34*

*p<.05. [a]Perceptual speed (Weschler, 1997); score was total number correct of 100 items. [b]Memory span (Weschler, 1997); score was total correct for the 14 sets of digits presented. [c]Verbal ability (Shipley, 1986); score was the total number correct from 40.

Participants were also tested for visual acuity using the Snellen eye-chart (Snellen, 1868). With the exception of one participant, all participants had a visual acuity of 20/40 for near vision (corrected or uncorrected). The participant who was an exception to the case had an uncorrected near vision of 20/60. However, her ability to see the instructions accurately was ascertained by requesting her to read aloud the questionnaire instructions. Additionally, her performance on the paper-pencil based ability tests was within 3 standard deviations from the mean of the older adult participants. Therefore, her data were included in the analysis.

Apparatus/Materials

The experiment was conducted using Dell Optiplex 760 computers running Microsoft Windows XP Professional. The system included a 17 inch monitor, configured to display 1280 x 1024 pixels. The software program was developed using E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA). The software program recorded participants' ratings for the stimuli presented across four different task contexts.

All on-screen textual instructions were displayed in the Calibri font in 18 font size. The software program presented all pictures on a computer monitor with the instructions. The pictures were shown with approximate pixel size of 240 X 168. Responses were collected using a standard QWERTY keyboard. Symbol key ` and numeral keys 2, 4, 6, and 8 were removed from the keyboard; keys 1, 3, 5, 7, and 9 were labeled as 1, 2, 3, 4, and 5 respectively to minimize inadvertent errors in key-pressing.

Humanoid Pictures

To manipulate levels of human-likeness, 4 robot faces and 4 human (2 male, 2 female) faces were selected. The human faces were chosen from the Montreal Set of Facial Displays of Emotions (MSFDE). All 4 faces had neutral expressions and were of White Caucasian actors.

The four robot faces corresponded to the humanoid robots: Pearl Nursebot, Nexi MDS (Mobile/Dextrous/Social), Nao, and Kobian. Pearl (Personal Robotic Assistants for the Elderly) is a research robot at Carnegie Mellon University, Nexi M.D.S. (Mobile/Dextrous/Social) is a research robot at Massachusetts Institute of Technology, Nao is a humanoid robot developed by a French company, Aldebaran Robotics, and Kobian is a Japanese humanoid robot. Humanoid robots vary in appearance characteristics in an unsystematic manner. Therefore, robot stimuli were chosen for this study to cover a range of humanoid facial appearance. However, all four robots had a pair of eyes and a resemblance of a mouth and/or a nose (Figure 2.1).

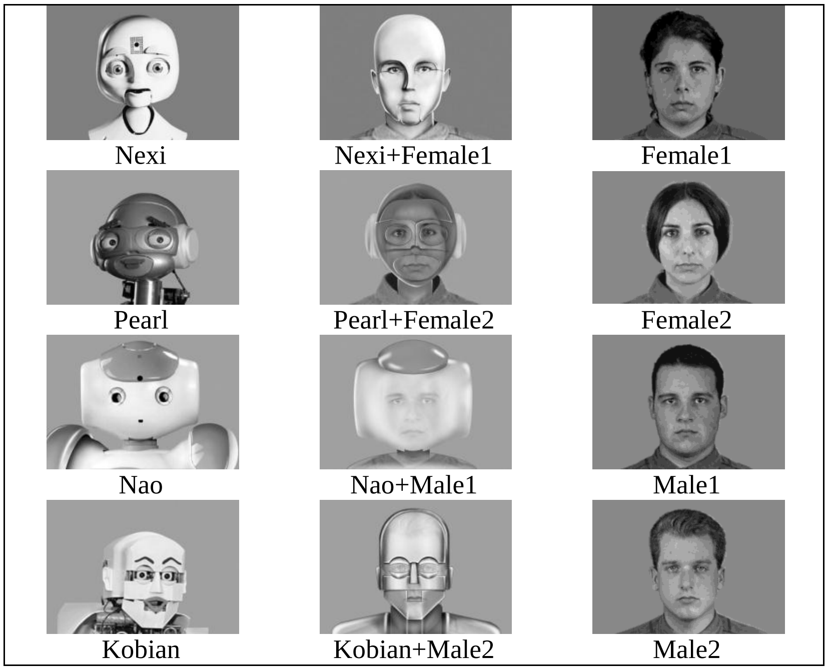


Figure 2.1. Pictures used in the study to represent different levels of humanoid appearance. (From left to right: robotic appearance, human-robot mixed appearance, and human appearance.

Each robot face was paired with a human face. For each robot-human pair, an appearance was created that lay between the human and the robot by morphing the two pictures. Thus, for every robot-human pair, the participants saw 3 appearances: completely human-like, between human-like and robot-like (mixed), and completely robot-like. In all, 4 such sets of face pictures were generated resulting in a total of 12 pictures (Figure 2.1). All pictures were converted to black and white and were cropped to be of the same size. The pictures were also printed on separated sheets and laminated to be presented to the participants during the interview.

Questionnaires

Four questionnaires were administered at different points in the study:

1. Demographics Questionnaire (Czaja et al., 2006; Appendix B) administered at the start of the study to collect demographics and general health information.
2. Robot Opinions Questionnaire (Smarr et al., 2012; Appendix C) is a robot-specific 12- item questionnaire modeled after technology acceptance scales (Davis, 1989). Participants respond to questions such as “My interaction with a robot would be clear and understandable”, “I would find a robot useful in my daily life”, and “Using a robot would make my daily life easier.” on a 7-point Likert scale (1 = Extremely unlikely, 4 = Neither unlikely or likely, 7 = Extremely likely).
3. Robot Facial Appearance Questionnaire (locally developed; Appendix D) assessed people’s opinions about the facial appearance of their imaginary robot. It consists of 15 items (e.g., “I would want my robot to have eyes”, “I would want my robot to look exactly like a human”, and “I would want my robot’s face to be unique”). Responses are marked on a 5 point Likert scale (1= strongly disagree, 3 = neither agree nor disagree, 5 = strongly agree).
4. Robot Familiarity and Use Questionnaire (Prakash et al., 2013, Smarr et al., 2012; Appendix E) requires participants to indicate their level of familiarity with 13 different kinds of robots on a 5-point scale (0 = not sure what it is; 4= have used or operated this frequently).
5. Assistance Preference Checklist (Prakash et al., 2013, Smarr et al., 2012; Appendix F) was the final questionnaire administered. The participants were asked to imagine they needed assistance in everyday life and then indicate preferences for human versus robot assistance with 58 home-based tasks, assuming the robot could perform those tasks to the level of a human. Assistance preference was indicated on a five-point scale (1=only a human, 3=no preference, 5=only a robot).

Design

The rating task was a 2 (age) X 3 (humanness) X 4 (task) split plot design where age group was a between subject factor; humanness and task were within subject factors. Age group consisted of two levels: younger adults and older adults. Humanness was the degree to which the robot face resembled a human face and comprised three levels: human appearance, mixed appearance, and robot appearance (Figure 2.1). Task had four levels: personal care (ADL), chores (IADL; low cognitive demand), decision-making (IADL; high cognitive demand), and social task (EADL).

Participants perceptions were assessed via four dependent measures (DVs): perceived usefulness, trust, likeability, and anxiety. Each DV consisted of a single item and the response was measured on a 5-point scale where 1= not at all, 3 = a fair amount, and 5 = a very much.

Procedure

The same experimenter administered the study to all the participants. Only one participant was present in the study at a time. At the start, participants signed an informed consent. Following this they filled out the demographics questionnaire and the robot opinions questionnaire. Next, they performed a rating task on a computer. Participants completed the ability tests between different sections of the rating task: The procedural flow of the study has been illustrated in Appendix G.

Rating Task

Participants were first given an overview of the rating task followed by a practice task. The first part of the practice task consisted of getting familiarized with the keys for selecting one’s responses. During the second part of the practice, participants were given a sample rating task wherein they were asked to imagine that they are at a department store and a robot is assisting them in finding items of their choice. Participants were then shown pictures of robots (different from the ones used in the study) and were asked to rate those on perceived usefulness, trust, likeability, and anxiety. At the end of practice, participants' doubts, if any, were clarified. Participants were also informed that the speed of at which they responded did not matter and that there were no right or wrong answers.

When the participant was ready to begin the rating task, the following set of instructions was displayed on the screen and was also read aloud (to minimize differences in participants' assumptions about robot capability, autonomy, and control):

"*Imagine that you need some assistance and that you have been given a robot to take home with you.*

- *The robot can perform tasks for you.*
- *You do not have to program the robot.*
- *You should assume that the robot can do what you want it to do.*
- *In this study we are focusing on the robot's face. Assume the robot's body to be consistent with the robot's facial appearance. The robot's body is such that it does not reduce its efficiency in performing a task."*

Next, participants were asked to imagine interacting with a robot in four different task scenarios (one scenario at a time with breaks in between the scenarios). The details of the task scenarios are delineated in Figure 2.2. In each scenario, participants were presented with a total of 12 face pictures (human, mixed, and robot; 4 of each; Figure 2.1). They were asked to imagine the robot to have the appearance as shown in every picture and then rate the robot in terms of how useful they would find it, how much they would trust it, how much they would like it, how anxious they would feel toward it in that task scenario. The participants provided their responses on a 5-point unipolar Likert-type scale where 1 = "not at all", 2 = a little, 3 = "somewhat", 4 = much, and 5 = "very much". The presentation of the pictures was randomized without any constraints. The order of

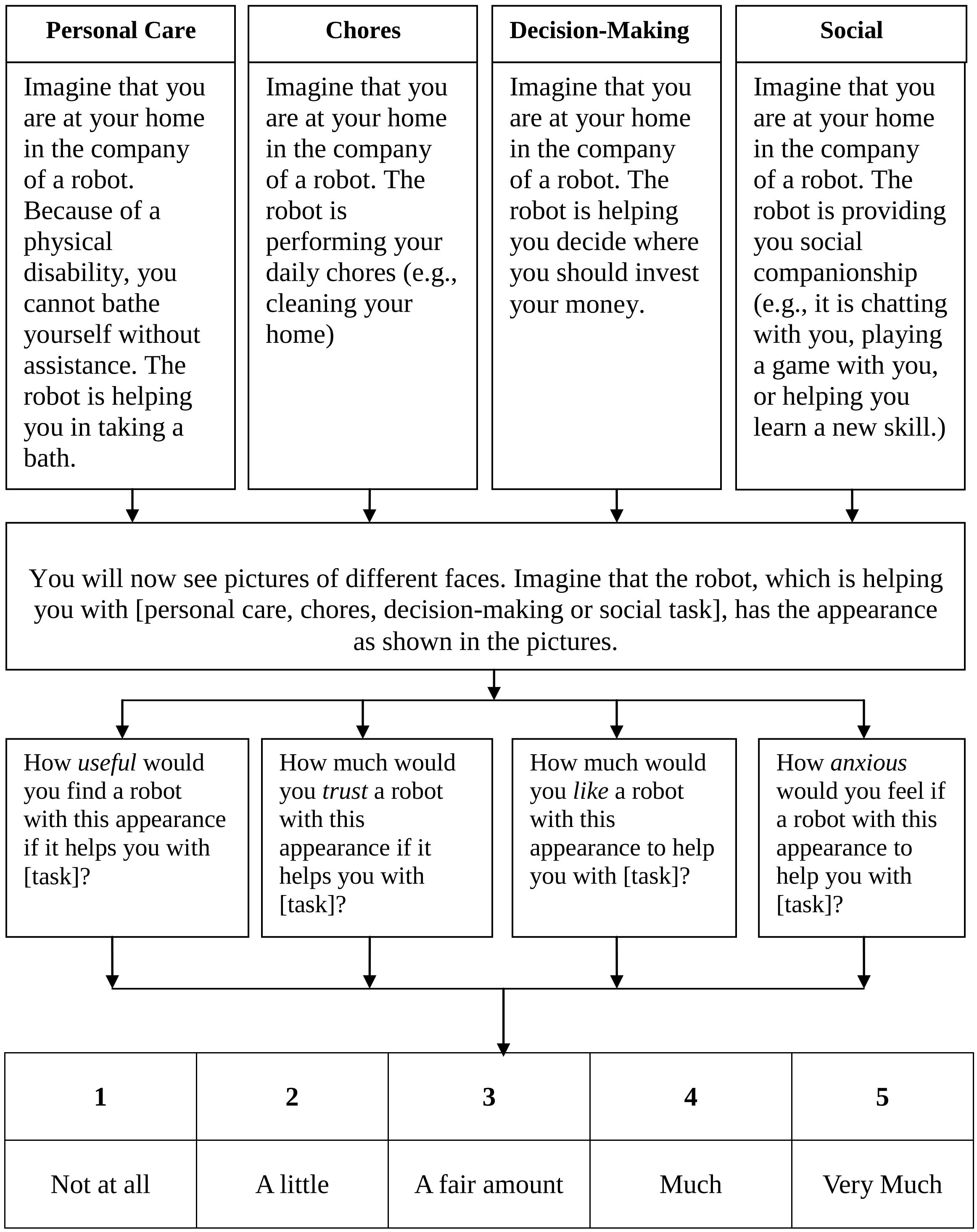


Figure 2.2. Schematic representation of the rating tasks

task contexts and rating measures were counterbalanced using a 4 X 4 partial Latin square design (Appendix H). Thus, in every task context, each picture was rated four times (once for each of the four DVs). Every participant completed a total of 192 ratings (= 12 pictures x 4 DVs x 4 tasks).

Appearance Preference Interview and Questionnaires

At the end of the rating task, participants were taken to another room where they were interviewed about their preferences for the robot facial appearance. The interview was audio-recorded. The interview script is presented in Appendix I. Participants were asked to imagine that they own a robot that stays with them in their home. The robot can assist them with all the tasks in the home that they imagined earlier during the rating task (i.e., it can bathe the person, perform daily chores, help in making investment decisions, and also provide them social companionship.)

Participants were presented with the two female human pictures (Figure 2.1) and were asked to select the one that they would prefer for their robot's appearance. Same was repeated with the male human pictures. The preferred female picture was then placed adjacent to the preferred male picture and participants were asked to decide which one they would prefer over the other. They were also asked to provide reasons for their selection. Similar selection tasks were performed for the mixed and the robotic appearance pictures.

Finally, the participant's most preferred human picture, the most preferred robot picture, and the most preferred mixed picture were placed together in front of the participant. The participant was asked to pick the most preferred appearance out of the three pictures, and also provide reasons for the choice. Next, participants were asked to think specifically about a personal care task (e.g., bathing) and were inquired if they would have a preference among the three faces if the robot helped them with that task. This was repeated for the other three categories of tasks: menial task (e.g., chores), social task (e.g., chatting with someone, playing a game with someone, or learning a new skill from someone), and decision-making task (e.g., deciding where to invest money). The order of the tasks was held constant for the interview. The rationale was that the chances of order effects were much less at this point as participants would have had enough exposure to every picture and to the four task contexts.

After the interview, participants filled out the robot facial appearance questionnaire, the robot familiarity and use questionnaire, and the assistance preference checklist. At the conclusion of the study, participants were debriefed and compensated for their time.

CHAPTER 3

RESULTS

Overview of Analysis

This section consists of three parts: results from the quantitative analysis of the rating task, results from the qualitative thematic analysis of the robot preference interview, and finally, results from the quantitative analyses of the questionnaire data (i.e., robot opinions; robot familiarity and use; robot facial appearance).

Results Part 1: The Rating Task

The rating task was a 2 (age) X 3 (humanness) X 4 (task) split plot design. Age is a grouping variable (between subjects) and humanness and task are within subject variables. There were four dependent variables (DVs): perceived usefulness (PU), trust, likeability, and anxiety. Participants' anxiety data are not used in the main analysis because the term "anxious" was not clearly understood by all participants and some participants were confused with the direction of the scale. Therefore, details about the anxiety data and their analysis are reported in Appendix J. The statistical test alpha was set at $p < .05$.

Do Perceptions Vary as a Function of Age, Humanness, and Task?

An omnibus MANOVA was performed to assess the effects of age, robot humanness, and task on people's *perceptions*, which is the underlying construct comprising PU, trust, and likeability. In other words, MANOVA was conducted to examine the three DVs simultaneously. Pillai's Trace (V) was used as it is more robust to the multivariate normality assumption than Wilk's Lamba (Λ; Fausset, Rogers, & Fisk, 2009). The statistical test alpha was set at $p < .05$. The results of the MANOVA are summarized in Table 3.1.

The analysis revealed a significant three-way age X humanness X task interaction. This suggests that the interaction effect of humanness X task on perceptions depended on the age. A two-way age X humanness interaction was also significant meaning that perceptions of the different levels of humanness depended on the age group. Additionally, there was a significant two-way interaction of age X task suggesting that the perceptions of robots across different tasks varied with the age. Finally the two-way interaction of humanness X task was also significant implying that for younger and older adults combined, perceptions for the different levels of humanness depended on the task.

Table 3.1

Summary of the Omnibus MANOVA Test

Effect	F-statistic	DF	p-value	partial η^2
Age X Humanness X Task	1.80*	18, 1116	0.02	0.03
Age X Humanness	2.45*	6, 246	0.03	0.06
Age X Task	2.06*	9, 558	0.03	0.03
Humanness X Task	4.08*	18, 1116	<0.001	0.06
Age	4.67*	3, 60	0.01	0.18
Humanness	6.54*	6, 246	<0.001	0.14
Task	5.56*	9, 558	<0.001	0.08

Note: DF is degrees of freedom

A main effect of age was found implying that overall, younger adults (M = 2.80, SD = 0.35) had more positive perceptions of robots compared to older adults (M = 2.69, SD = 0.66). Additionally, there was a main effect of humanness suggesting that for younger and older adults combined, perceptions varied as a function of robot humanness.

Moreover, a main effect of task was found suggesting that for younger and older adults considered together, perceptions varied across task.

To further investigate the statistically significant omnibus MANOVA, three separate ANOVAs were conducted on each of the three DVs (Fausset et al., 2004). The type I error rate was Bonferroni corrected (0.05/3). Therefore, the critical alpha level was set at $p < 0.0167$ for all further analyses. Huynh-Feldt corrections were applied where sphericity assumptions were violated.

Before proceeding to the univariate analyses on the DVs, it is to be noted that the three DVs were highly correlated, especially for older adults (see Table 3.2). Moreover, PU and trust were more strongly correlated than likeability and trust, and likeability and PU. Such a high correlation between PU and trust implies that two DVs might be measuring the same underlying construct.

Table 3.2

Correlation Matrix of Perceived Usefulness (PU), Trust, and Likeability

Age	Variable	SD	PU	Trust	Anxiety
Younger Adults	PU	(.42)			
	Trust	(.37)	.88**		
	Likeability	(.36)	.66**	.68**	
Older Adults	PU	(.63)			
	Trust	(.72)	.96**		
	Likeability	(.69)	.89**	.87**	

**Correlation is significant at the 0.01 level (2-tailed)

How do Perceptions of Usefulness vary as a Function of Age, Humanness and Task?

As observed in Figure 3.1, on average younger participants' ratings of PU were around 3 (= a fair amount) although they evaluated mixed appearance much less favorably for the social task. The older age group's average PU ratings were around 3 only for the human appearance. Apparently, they evaluated other appearances less favorably on PU for all the tasks except chores. Both the age groups seemed to consider robots of all levels of humanness more useful for assistance with chores (with average ratings equal to or slightly above 3) than with other tasks.

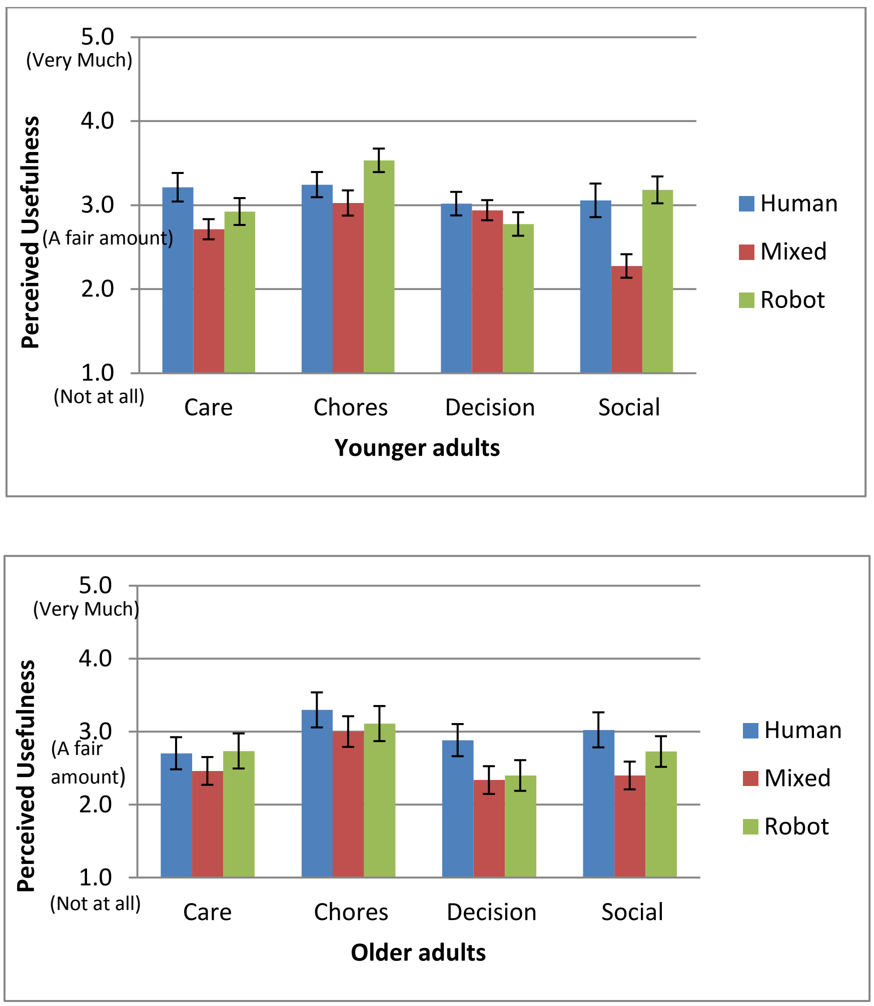


Figure 3.1. Mean PU ratings by age, humanness, and task. Error bars represent standard error of the mean.

An age X humanness X task ANOVA yielded a significant three-way interaction (F (4.1, 254.37) = 3.14, p = 0.01, η_p^2 = 0.05). This implied that the interactive effect of humanness and task on PU differed for younger and older adults (Figure 3.1). Therefore, humanness X task repeated measures ANOVA was performed separately for the younger and older adult age groups. Humanness X task interaction was significant for the younger adults (F (4.61, 142.85) = 6.10, p < 0.001, η_p^2 = 0.16) but *not* significant for the older adults (F (2.58, 80.10) = 1.67, p = 0.19, η_p^2 = 0.05).

Pair-wise comparisons of humanness X task for the younger adults revealed that this age group considered mixed appearance to be least useful for the social task as compared to chores, personal care, and decision-making. Moreover, younger adults perceived robot appearance as more useful for chores than for personal care and decision-making tasks. For the younger adult group, paired t-tests were also conducted conditioned on the task. The analyses revealed that for social task, younger adults perceived the mixed appearance as least useful compared to the other two appearances. For personal care task, they perceived human appearance as more useful than the mixed appearance whereas for chores, robot appearance was perceived more useful than the mixed appearance.

<u>How Do Interactions of Age and Humanness, Age and Task, and Humanness and Task Affect Perceptions of Usefulness?</u>

Of the three possible two-way interactions, only humanness X task interaction (F (6, 372) = 4.79, p = 0.001, η_p^2 = 0.07) was found significant (p < 0.0167). Thus, for younger and older adults combined, perceptions of usefulness for the levels of humanness varied across the tasks. To investigate the interaction further, post-hoc comparisons (paired-tests at p < 0.0167) were performed on the means by conditioning on humanness and task separately.

When conditioned on humanness, paired t-tests revealed that the mixed appearance was perceived as more useful for chores than for the other three tasks. Moreover, the mixed appearance was perceived least useful for the social task. The robot appearance was also perceived most useful for chores; it was perceived less useful for decision-making (compared to chores and social task). When conditioned on the task, paired-tests revealed that for the social task, mixed appearance was perceived less useful than human and robot appearance. For chores, robot appearance was perceived more useful than mixed appearance whereas for personal-care task, human appearance was perceived more useful than mixed appearance.

Summary of the Interaction Effects of Age, Humanness, and Task on Perceived Usefulness

The younger adults' perceptions of usefulness for different levels of humanness varied with the task. They perceived less use for mixed appearance for social task relative to other tasks. Moreover, they perceived more use of robot appearance for chores than for personal care and decision-making. Older adults' perceptions of different levels of humanness did not vary significantly across tasks.

With the age-groups combined, and comparing across the four tasks, mixed appearance was perceived as most useful for chores and least useful for the social task. Robot appearance was also perceived most useful for chores, but less useful for decision-making. Perceived usefulness for human-appearance did not reveal distinct patterns across different tasks.

<u>How do Age, Humanness, and Task Separately Affect Perceptions of Usefulness?</u>

Task had a significant main effect on perceived usefulness (F (3, 186) = 11.81, p < 0.001, $\eta_p^2 = 0.16$). Paired t-tests were performed on the marginal means for task (averaged across the three levels of humanness). The results suggest that compared to social, personal-care, and decision-making tasks, robots are perceived most useful for chores (Figure 3.2). There was no main effect of age on PU (F (1, 62) = 2.69, p = 0.11, η_p^2 = 0.04).

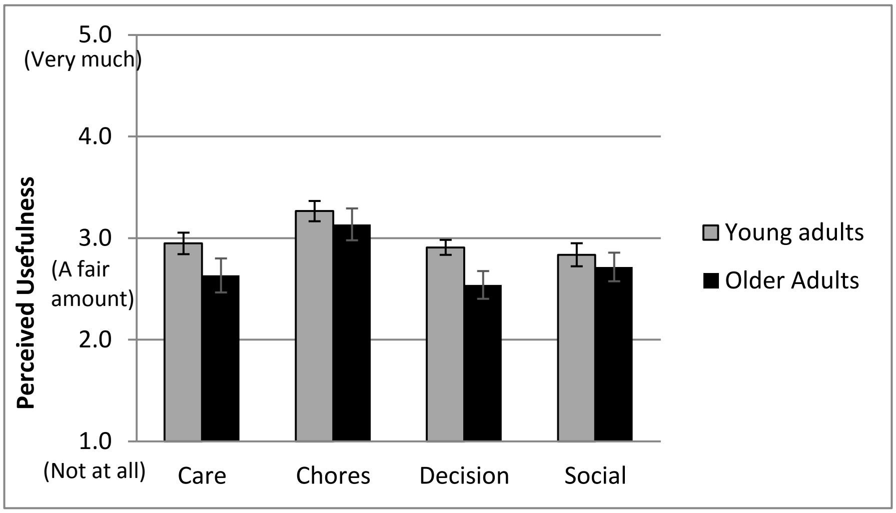


Figure 3.2. Mean PU ratings by task. Error bars represent standard error of the mean.

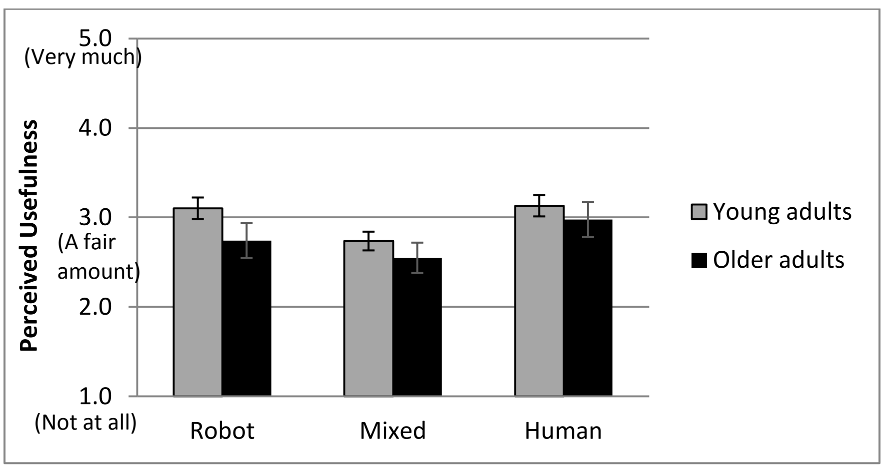


Figure 3.3. Mean PU ratings by humanness. Error bars represent standard error of the mean.

Although the marginal mean for mixed appearance (M = 2.64, SD = 0.80) was less than that for human (M = 3.05, SD = 0.92) and robot (M = 2.92, SD = 0.93) appearances, a significant main effect of humanness was not obtained (F (22.56, 336.59)

= 4.16, $p = 0.03$, $\eta_p^2 = 0.06$). However, perceptions of usefulness were not uniform within each level of humanness (i.e., human, mixed, and robot appearances).

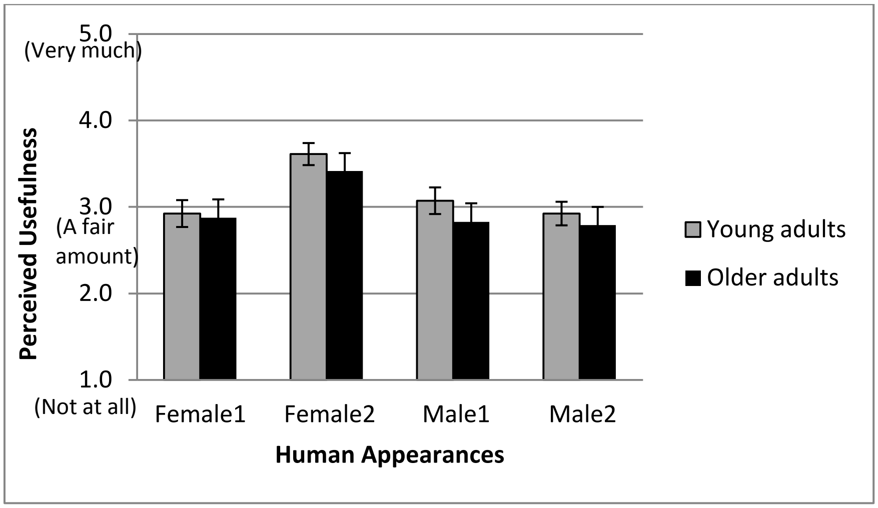


Figure 3.4. Mean PU ratings for different human appearances. Error bars represent standard error of the mean.

An age (2) X human appearance (4) split plot ANOVA revealed a significant main effect of human appearance (F (2.88, 178.47) = 24.63, $p < 0.001$, $\eta_p^2 = 0.28$) implying that participants' perceptions of usefulness varied across different human appearances. The effect of age and the interaction of age X human appearance were not significant. Paired t-tests were performed to analyze further the main effect of human appearance. Perceptions of usefulness were significantly higher ($p < 0.001$) for Female2 than for the other three human faces (see Figures 2.1, 3.4). In fact, this was the only appearance (of all the human, mixed, and robot appearances) for which older adults' mean PU was above 3 (3 = a slight amount).

An age (2) X mixed appearance (4) split plot ANOVA also produced a significant main effect of mixed appearance (F (3, 186) = 14.20, $p < 0.001$, $\eta_p^2 = 0.19$) as well as a significant interaction effect (F (3, 186) = 5.60, $p = 0.001$, $\eta_p^2 = 0.08$) Post-hoc comparisons revealed that younger adults perceived Nexi+Female1 as most useful compared to the other mixed appearances ($p < 0.001$; see Figures 2.1, 3.5). Older adults perceived Nao+Male1 as less useful than Nexi+Female1 ($p < 0.01$) and Pearl+Female2 ($p < 0.001$).

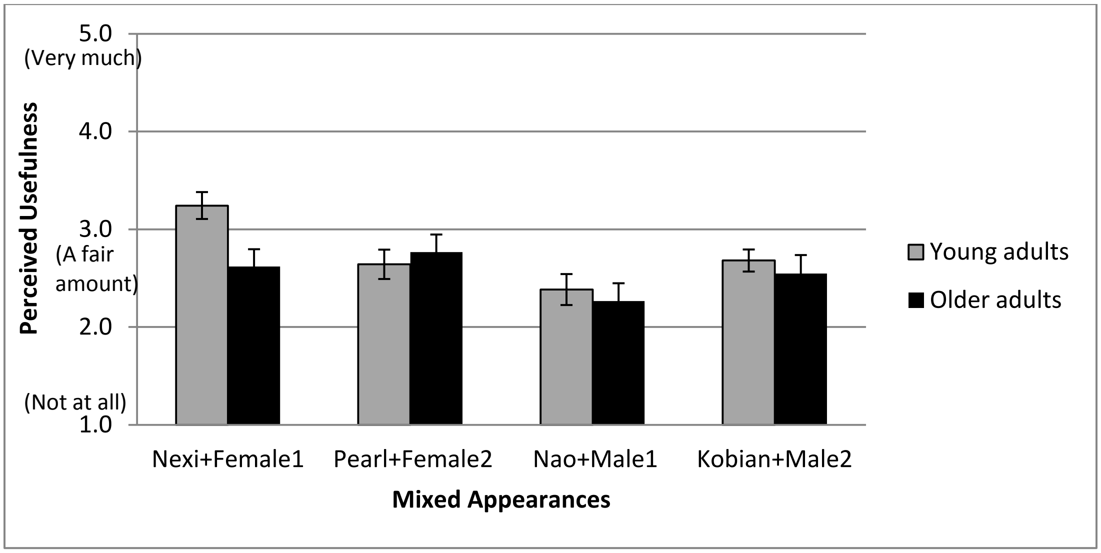


Figure 3.5. Mean PU ratings for different mixed appearances. Error bars represent standard error of the mean.

Finally, an age (2) X robot appearance (4) split plot ANOVA was also conducted. A significant main effect of robot appearance was found (F (2.7, 167.51) = 16.18, $p < 0.001$, $\eta_p^2 = 0.21$). Additionally, the interaction of age and robot appearance was also significant (F (2.70, 167.51) = 7.34, $p < 0.001$, $\eta_p^2 = 0.11$). Post-hoc paired t-tests revealed that younger adults perceived Kobian to be least useful compared to the other robots ($p < 0.001$; See Figures 2.1, 3.6). Older adults perceived Kobian to be less useful than Nexi and Pearl. Moreover, older adults perceived Nao as less useful than Nexi.

Of all the 12 appearances, older adults' mean PU rating was highest and above the mark of 3 (where 3 = a fair amount) for Female2 (M = 3.44, SD = 1.15) followed by Nexi (M = 3.02, SD = 1.25). Younger adults' mean PU was also highest for Female2 (M =

3.61, *SD* = 0.72). Younger adults' mean PU was above 3 for six of the twelve appearances (i.e., Female2, Male1, Nexi+Female1, Nexi, Pearl, and Nao).

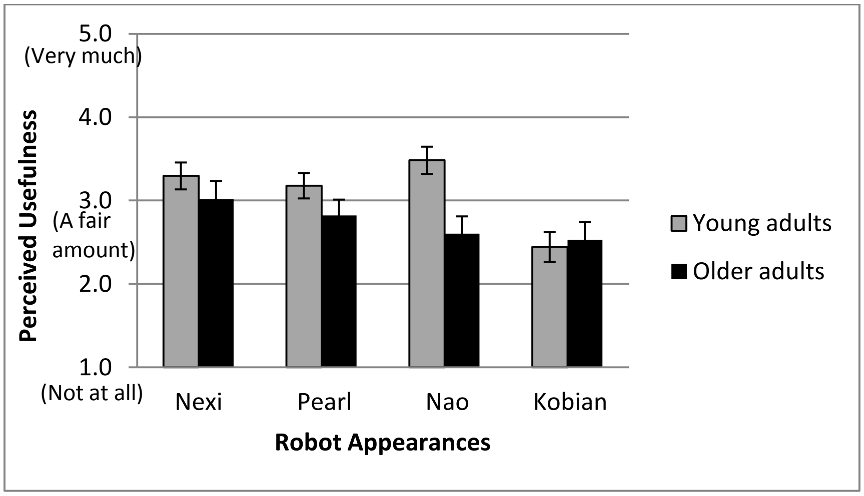


Figure 3.6. Mean PU ratings for different robot appearances. Error bars represent standard error of the mean.

Summary of the Main Effects of Age, Humanness, and Task on Perceived Usefulness

Only task had a significant main effect on PU such that participants considered robots to be most useful for chores than for other tasks. Although humanness of the robot did not have a significant effect on PU, participants did not evaluate all the human-appearances as equally useful. One of the female robots (Female2) was perceived more useful than all the other humans. Similarly, significant differences in PU were observed *within* the mixed appearance and robot appearance categories. Younger adults perceived Nexi+Female1 mixed appearance as most useful of the mixed appearances. Older adults perceived Nexi+Female1 and Pearl+Female1 appearances as more useful than Nao+Male1. In the robot appearance category, younger adults perceived Kobian as least useful of the other robots; older adults also perceived Kobian as less useful, significantly so when compared to Nexi and Pearl.

Overall Summary of the Effects of Age, Humanness, and Task on Perceived Usefulness

Perceptions of usefulness for different levels of humanness depended on the task and this relationship was further moderated by the age. Younger adults perceived mixed appearance as least useful for the social task with respect to other tasks. They perceived robot appearance to be more useful for chores than for personal care and decision making.

With both the age groups combined, mixed appearance was again perceived as least useful for the social task. Of all the tasks, mixed appearance was perceived most useful for chores. For the two age groups considered together, and comparing across different tasks, robot appearance was also perceived as most useful for chores but less useful for decision-making.

Participants perceived robots to be most useful for performing chores over the other tasks. There was no main effect of age or humanness (at $p < 0.0167$; although marginal means of PU for mixed appearance were less than that for human and robot appearances). Additionally, participants' perceptions of usefulness varied *within* the three levels of humanness (human, mixed, and robot).

How do Perceptions of Trust vary as a Function of Age, Humanness and Task?

The trends observed in Figure 3.7 indicate that in general both younger and older adults' ratings of trust were between 2(= a little) and 3 (= a fair amount). However, younger adults' mean ratings of trust exceeded the mark of 3 for robot appearance for chores and social task. For both the age groups, trust ratings appeared to be lower for mixed appearance compared to the other appearances. The only exception to this trend was noted for the decision-making task where trust in mixed appearance and robot appearance seemed comparable.

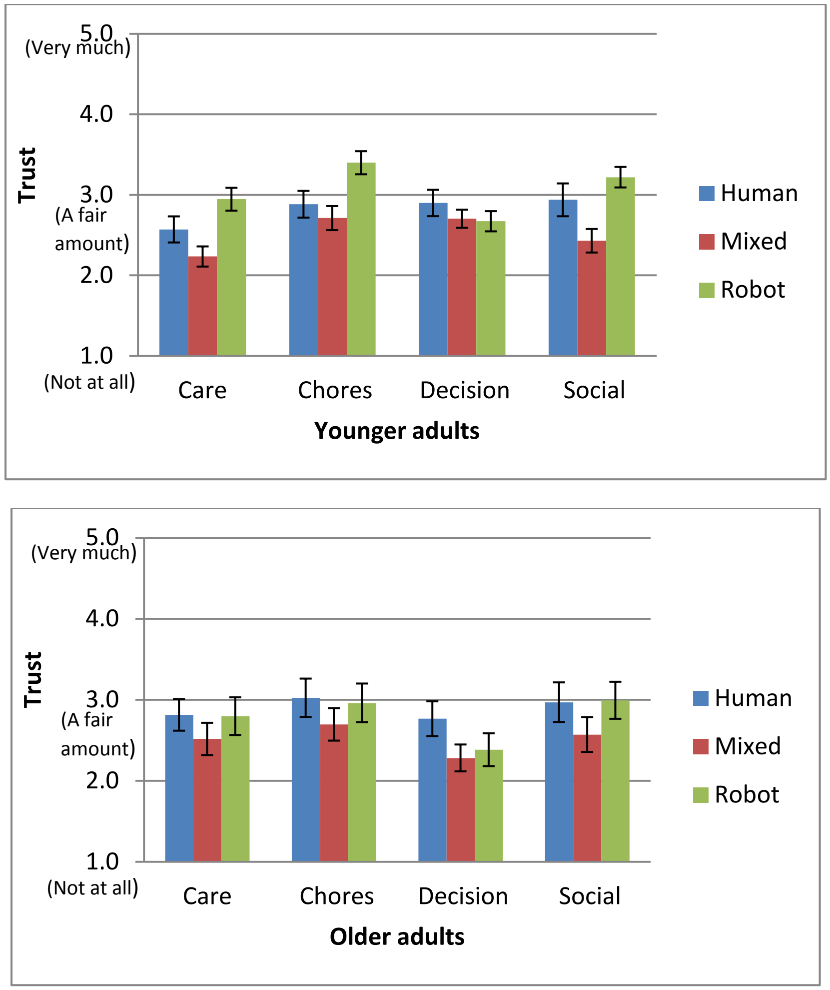


Figure 3.7. Mean trust ratings by age, humanness, and task. Error bars represent standard error of the mean.

How Do Interactions of Age and Humanness, Age and Task, and Humanness and Task Affect Perceptions of Trust?

A univariate analysis of age X humanness X task did not reveal a significant three-way interaction (F (3.86, 239.24) = 1.26, p = 0.28, η_p^2 = 0.02). Two-way interactions of age X humanness (F (4.24, 336.59) = 0.78, p = 0.42, η_p^2 = 0.01) and age X task (F (4.15, 146.89) = 2.09, p = 0.10, η_p^2 = 0.03) were also not significant. However, a significant two-way interaction of humanness X task (F (8.43, 239.24) = 4.77, p = 0.001,

$\eta_p^2 = 0.07$) was obtained. This meant that participants trust across different levels of humanness depended on the task.

Post-hoc comparisons (paired-tests at $p < 0.0167$) were performed to examine the humanness X task interaction. When conditioned on the humanness, it was found that participants would trust a mixed appearance more when the robot performed chores than when it performed a personal care task ($p < 0.001$; Figure 3.7). Additionally, they would trust a robot appearance less for a decision-making task compared to personal care ($p = 0.001$), chores ($p < 0.001$), and social task ($p < 0.001$). Moreover, they would trust a robot appearance less for personal care than for chores ($p = 0.01$).

When conditioned on the task, paired t-tests indicated that a robot appearance would be trusted more than a mixed appearance for personal care ($p = 0.001$), chores ($p < 0.001$), and social task ($p < 0.001$), but not for decision-making ($p = 0.75$). Human appearance was trusted significantly more than the mixed appearance only for the social task ($p = 0.01$).

Summary of the Interaction Effects of Age, Humanness, and Task on Trust

A robot appearance was least trusted for decision-making task than the other tasks. This appearance was trusted more for chores than for personal care and decision making. The mixed appearance would be trusted more for chores when compared to personal care (but not when compared to decision-making, unlike robot appearance). Across all tasks but decision-making, robot appearance is trusted more than mixed appearance. Human appearance was trusted more than mixed appearance for social task.

<u>How do Age, Humanness, and Task Separately Affect Perceptions of Trust?</u>

A significant main effect of task was observed (F (3, 186) = 7.39, $p < 0.001$, $\eta_p^2 = 0.11$). Paired t- tests on the marginal means (averaged across humanness) showed that robots would be trusted more for chores than for personal care and for decision-making

(p's < 0.001; Figure 3.8). Additionally, robots would be trusted more for social task than for personal care and decision making (p = 0.016 and 0.013 respectively).

The main effect of age was not significant (F (1, 62) = 0.23, p =0.63, $\eta_p^2 = 0.00$). The main effect of humanness was also not significant at $p < 0.0167$ (F (1.43, 88.54) = 4.42, p = 0.02, $\eta_p^2 = 0.07$) but was significant at p = 0.05 level. As was noted for perceived usefulness, the comparison of marginal means indicated less trust for the mixed appearance than for human and robot appearances (Figure 3.9).

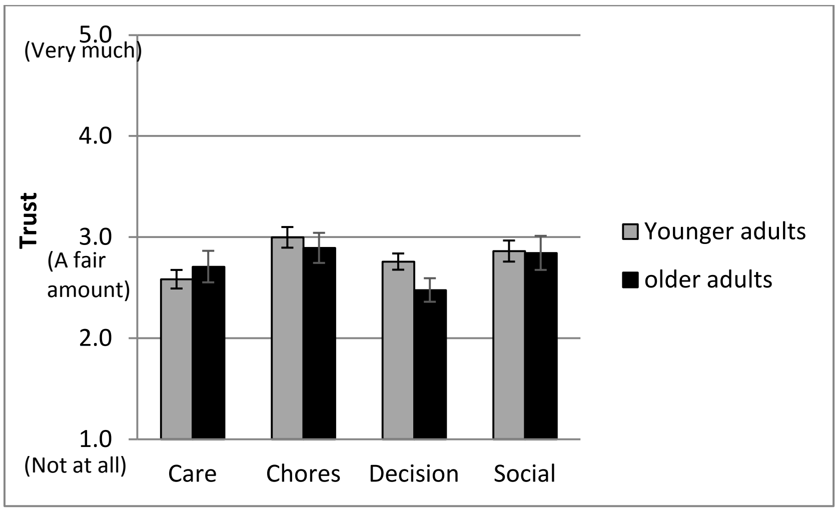


Figure 3.8. Mean trust ratings by task. Error bars represent standard error of the mean.

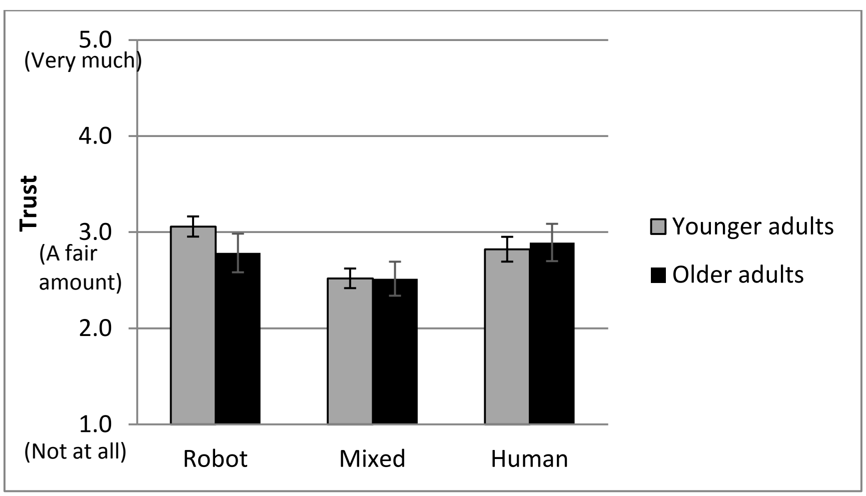


Figure 3.9. Mean trust ratings by humanness. Error bars represent standard error of the mean.

Three split-plot ANOVAs were conducted to assess if participants' trust varied across the two age groups and within the levels of humanness (i.e., across the four human, mixed, and robot appearances respectively). The age (2) X human appearance (4) split-plot ANOVA yielded a main effect of human appearance (F (3, 186) = 30.20, $p <$ 0.001, $\eta_p^2 = 0.33$). This implied that participants' trust in a human-looking robot would vary across different human appearances. Post-hoc comparisons revealed that Female2 was trusted significantly more than the other three humans (p < 0.001; Figures 2.1, 3.10).

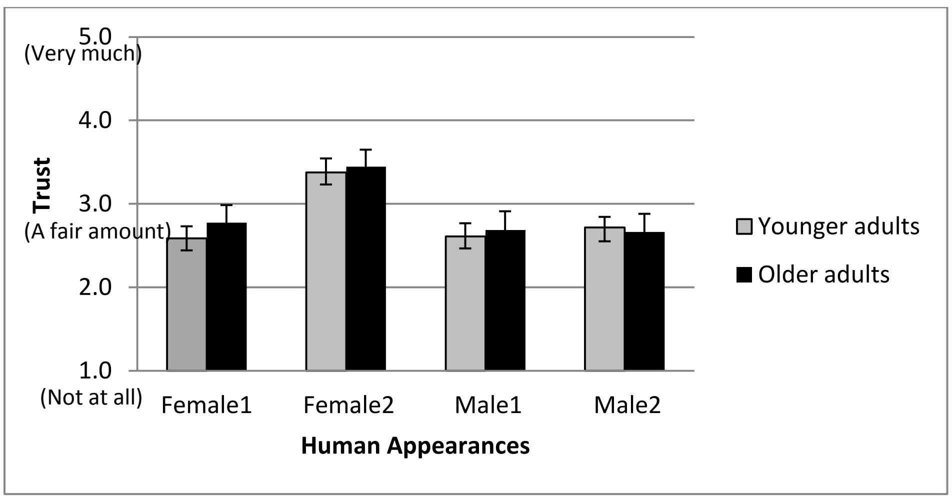


Figure 3.10. Mean trust ratings for different human appearances. Error bars represent standard error of the mean.

The age (2) X mixed appearance (4) split-plot ANOVA also yielded a main effect of mixed appearance (F (2.82, 174.61) = 11.92, p < 0.001, $\eta_p^2 = 0.16$) meaning that participants' trust in mixed appearance robots would depend on the kind of mixed appearance. Post-hoc paired t-tests revealed that Nao+Male1 was trusted less than the Nexi+Female1 and Pearl+Female2 (p < 0.001). Moreover, Kobian+Male2 was also trusted less than Nao+Male1 (p = 0.001) and Pearl+Female2 (marginal significance; p =

0.02). In general, this might suggest that the robots morphed on female faces were trusted more than the robots morphed on male faces (see Figures 2.1, 3.11). The main-effect of age and the interaction of age X mixed appearance were not significant (at $p < 0.0167$)

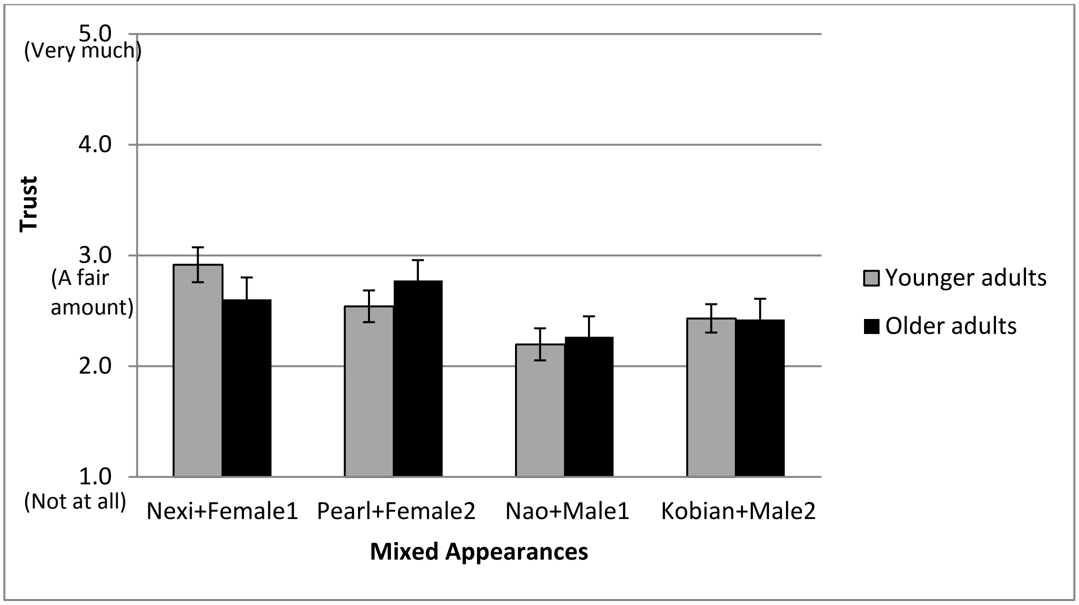


Figure 3.11. Mean trust ratings for different mixed appearances. Error bars represent standard error of the mean.

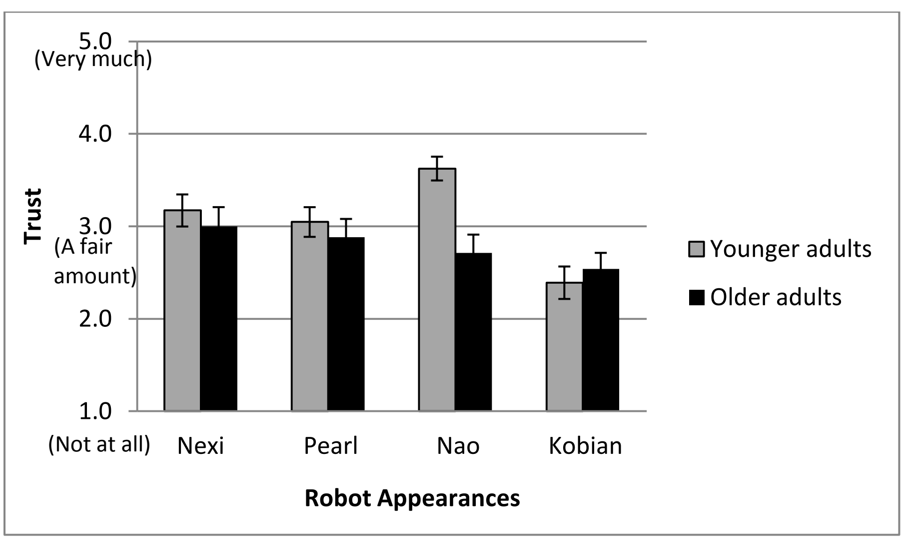


Figure 3.12. Mean trust ratings for different robot appearances. Error bars represent standard error of the mean.

Finally, an age (2) X robot appearance (4) split plot ANOVA produced significant main effect of robot appearance (F (3, 186) = 15.71, $p < 0.001$, $\eta_p^2 = 0.20$) and significant age X robot appearance interaction (F (3, 186) = 8.04, $p < 0.001$, $\eta_p^2 = 0.11$). There was no main effect of age. Post-hoc analysis of the significant effects revealed that for younger and older adults combined, Kobian is the least trusted robot ($p < 0.001$). When younger adults' data was analyzed separately, it was found that this age group trusted Kobian the least (p's < = 0.001). They also trusted Nao more than the other robots. Similar to younger adults, the older adults trusted Nexi and Pearl more than Kobian; however, unlike younger adults, they seemed to trust Nao less than they did Nexi and Pearl (see Figures 2.1, 3.12).

Of all the 12 appearances (human, robot, mixed), older adults' mean trust rating exceeded the mark of 3 (where 3 = a fair amount) only for Female2 (M = 3.44, SD = 1.15). Younger adults' mean trust rating was above 3 for Female2 (M = 3.37, SD = 0.95) and for three of the four robot appearances (Nexi, Pearl, and Nao). Younger adults' mean trust was highest for Nao (M = 3.62, SD = 0.92).

Summary of the Main Effects of Age, Humanness, and Task on Trust

Only task had a significant main effect on trust such that robots were trusted more for help with chores than with personal care and decision making. The main effect of age was not significant. The main effect of humanness was not significant at $p < 0.0167$ but was significant at a less conservative criterion of $p < 0.05$. The comparison of means indicated that overall a mixed appearance might be trusted less than a human or robot appearance. Additionally, perceptions of trust varied within every level of humanness.

In the human-appearance category, participants trusted Female1 the most. The patterns of trust were more complex for the mixed appearance and robot appearance. In the mixed appearance category, the general observation was that robots morphed on the female faces were trusted more than the robots morphed on the male faces. At this level of analysis, it is not possible to infer what led to the difference in trust across the four mixed appearances: the aesthetics of the constituent faces (human and robot) that formed the mixed appearance, or the resulting mixed appearance itself.

In the robot appearance category, younger and older adults showed different patterns of trust. While Kobian did seem to be the least trusted robot for both the age groups, younger adults trusted Nao the most whereas older adults had more trust in Nexi and Pearl.

Overall Summary of the Effects of Age, Humanness, and Task on Trust

Participants' trust in robots of varying levels of humanness depended on the task. The mixed appearance was trusted more for chores than for personal care. The robot appearance was trusted more for chores than for personal care and decision-making. In fact, the robot appearance was trusted *least* for decision-making compared to other tasks. Additionally, of the four tasks considered, decision making is the only one for which robot appearance is *not* trusted more than mixed appearance.

For all appearances considered together, a robot would be trusted more for assistance with chores than with personal-care and decision-making tasks. For all tasks considered together, a mixed appearance is likely to be trusted less than the human and robot appearances. However, all mixed appearances would not be trusted uniformly. Same holds for human and robot appearances.

How do Perceptions of Likeability Vary as a Function of Age, Humanness, and Task?

Participants' mean likeability ratings were generally between 2 (a little) and 3 (= a fair amount) although younger adults' likeability for robot appearance exceeded 3 for chores and social task (Figure 3.13). Younger adults seemed to like robot appearance more than human and mixed appearances for all tasks except decision-making. Older adults' likeability ratings for human and robot appearances seemed comparable for all tasks except decision-making for which likeability for robot appearance dropped considerably.

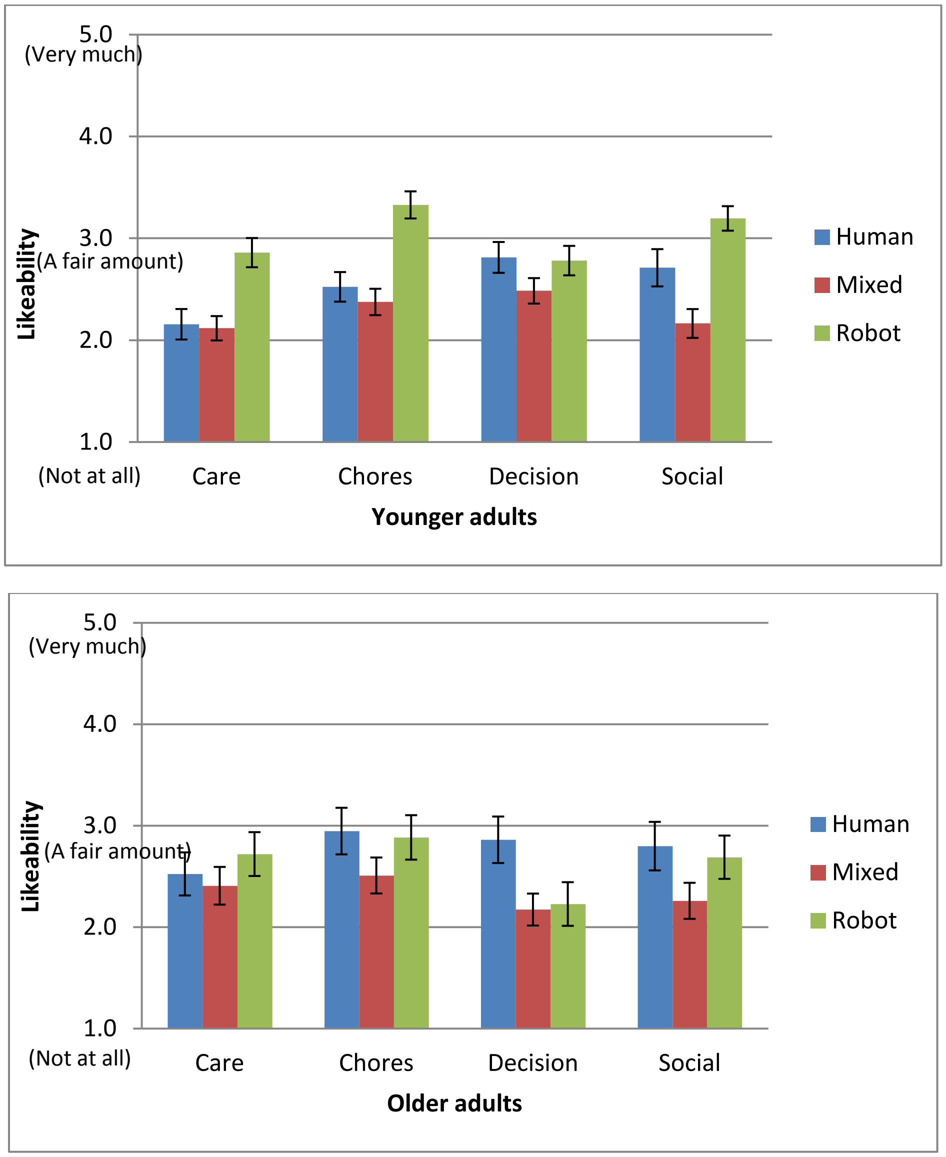


Figure 3.13. Mean likeability ratings by age, humanness, and task. Error bars represent standard error of the mean.

How do Interactions of Age and Humanness, Age and Task, and Humanness and Task Affect Perceptions of Likeability?

The univariate analysis of age X humanness X task on likeability did not yield a significant three-way interaction (F (4.74, 293.91) = 0.73, p = 0.59, $\eta_p^2 = 0.01$). The two-way interactions of age X humanness (F (1.48, 91.97) = 2.81, p = 0.08, $\eta_p^2 = 0.04$) and age X task (F (2.45, 151.69) = 2.47, p =0.08, $\eta_p^2 = 0.04$) were also non-significant. However, the humanness X task interaction was significant (F (4.74, 293.91) = 8.27, p < 0.001, $\eta_p^2 = 0.12$).

The humanness X task interaction was investigated further via post-hoc comparisons. Paired-tests were performed by conditioning on humanness and task separately. Conditioning on appearance and comparing across tasks, it was found that human appearance would be least liked for personal care (p's < 0.0167; Figure 3.13). Mixed appearance would be liked less for social task than for chores (p < 0.0167). Robot appearance would be liked the least for decision-making compared to personal care (marginally significant; p = 0.02), chores (p < 0.001) and social task (p < 0.001).

When conditioned on task, and compared across humanness, paired t-tests revealed that with the exception of the decision-making task, robot appearance would be liked more than the mixed appearance (p < 0.001). The human appearance would be liked more than the mixed appearance for decision-making and social tasks (p < 0.0167). Additionally, robot appearance would be liked more than the human appearance for assistance with a personal care task (p = 0.01).

Summary of the Interaction Effects of Age, Humanness, and Task on Likeability

As was seen for the other DVs, humanness and task had a significant interaction effect on likeability. Thus, the participants' likeability for the different levels of human appearance was moderated by the task. Human appearance was liked the least for assistance with a personal-care task, while robot appearance was least liked for a decision-making task. Likeability for mixed appearance was less for social task than for chores.

With the exception of the decision-making task, robot appearance was liked more than the mixed appearance for all tasks. Robot appearance was also liked more than human appearance for assistance with personal care. Human appearance was liked more than the mixed appearance for decision-making and social tasks.

How do Age, Humanness, and Task Separately Affect Perceptions of Likeability?

The main-effect of age (F (1.48, 91.97) = 7.32, p = 0.003, η_p^2 = 0.10) was not significant, but significant main effects were observed for task (F (2.45, 151.69) = 4.27, p = 0.01, η_p^2 = 0.06) and humanness (F (1.48, 91.97) = 7.32, p = 0.003, η_p^2 = 0.10). This meant that participants' likeability depended on the task the robot would assist with, as well as on the humanness of the robot face. Paired t-tests revealed that robots were liked more for assistance with chores than with personal-care and decision-making tasks (p's < 0.0167; Figure 3.14). Post-hoc comparisons also revealed that mixed appearance was liked less than human appearance (p = 0.01) and robot appearance (p < 0.001; Figure 3.15).

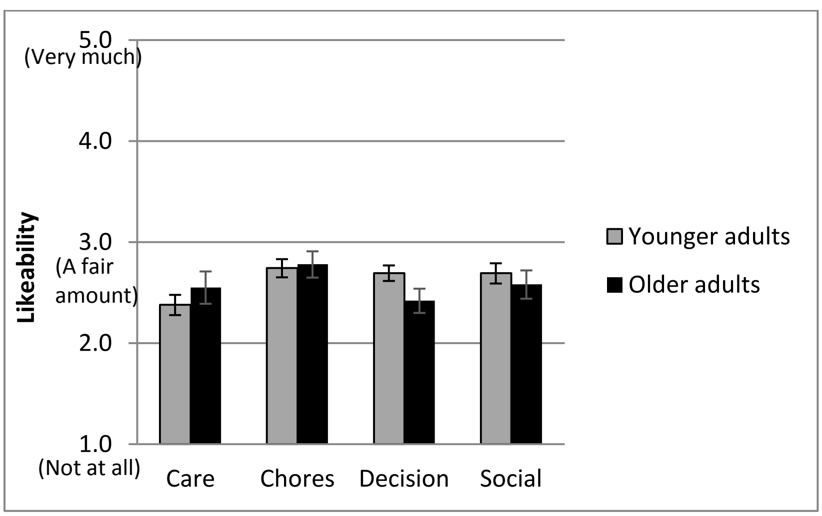


Figure 3.14. Mean likeability ratings by task. Error bars represent standard error of mean.

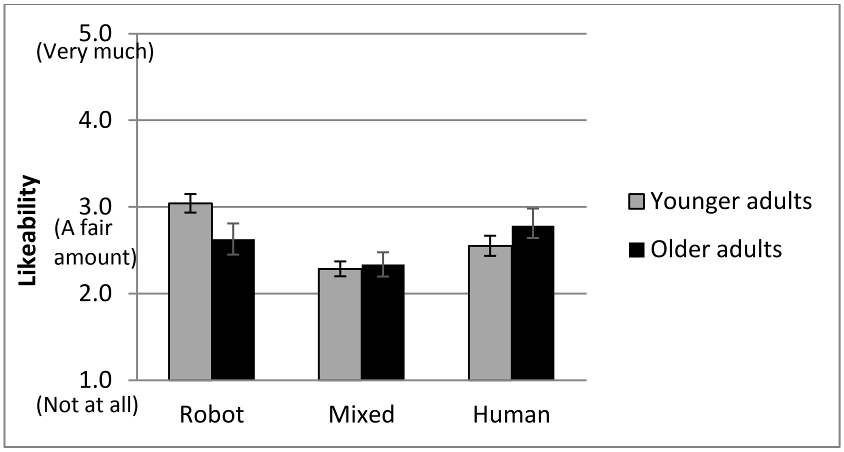


Figure 3.15. Mean likeability ratings by humanness. Error bars represent standard error of mean.

To investigate if younger and older adults' likeability for robots varied across the four human appearances, an age (2) X human appearance (4) split plot ANOVA were conducted. The analysis revealed a significant main effect of human appearance (F (3, 186) = 26.55, $p < 0.001$, $\eta_p^2 = 0.30$). Age neither had a significant main effect, nor a significant interaction effect with human appearance. Post-hoc analysis of the main effect of human appearance showed that participants' likeability was significantly higher for female2 compared to other human appearances ($p < 0.001$; see Figures 2.1, 3.16).

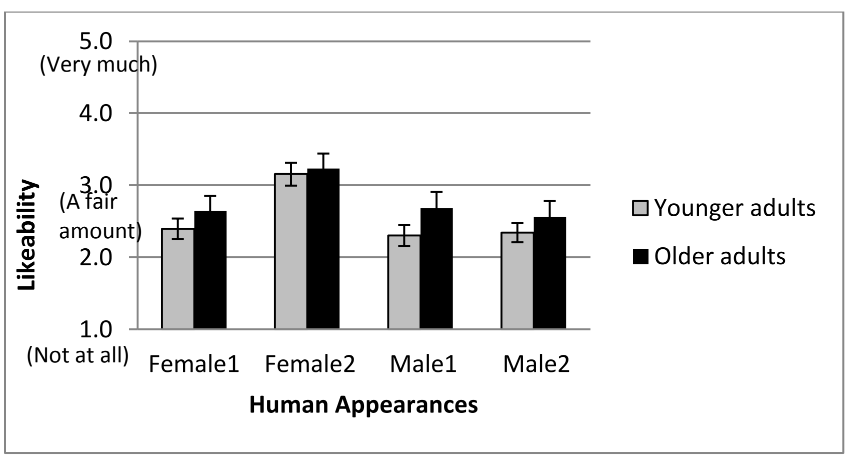


Figure 3.16. Mean likeability ratings for different human appearances. Error bars represent standard error of the mean.

Similarly, an age (2) X mixed appearance (4) split plot ANOVA was conducted to assess if the two age groups' mean likeability varied across the four mixed appearances. The analysis resulted in a significant main effect of mixed appearance (F (3, 186) = 13.78, $p < 0.001$, $\eta_p^2 = 0.18$) and a significant interaction of age X mixed appearance (F (3, 186) = 6.29, $p < 0.001$, $\eta_p^2 = 0.09$). Post-hoc comparisons showed that collapsed across both the age groups, Nao+Male1 was the least liked mixed appearance (p's < = 0.001) followed by Kobian+Male2. When paired t-tests were conducted separately for the two age groups, it was found that younger adults liked Nexi+Female1 the most (for comparison with Pearl+Female2, only marginal significance observed; $p = 0.02$). Younger adults' least liked mixed appearance was Nao+Male1 (marginal significance at p < 0.05). Older adults liked Pearl+Female1 the most and Nao+Male1 the least (all p's < 0.0167; see Figures 2.1, 3.17).

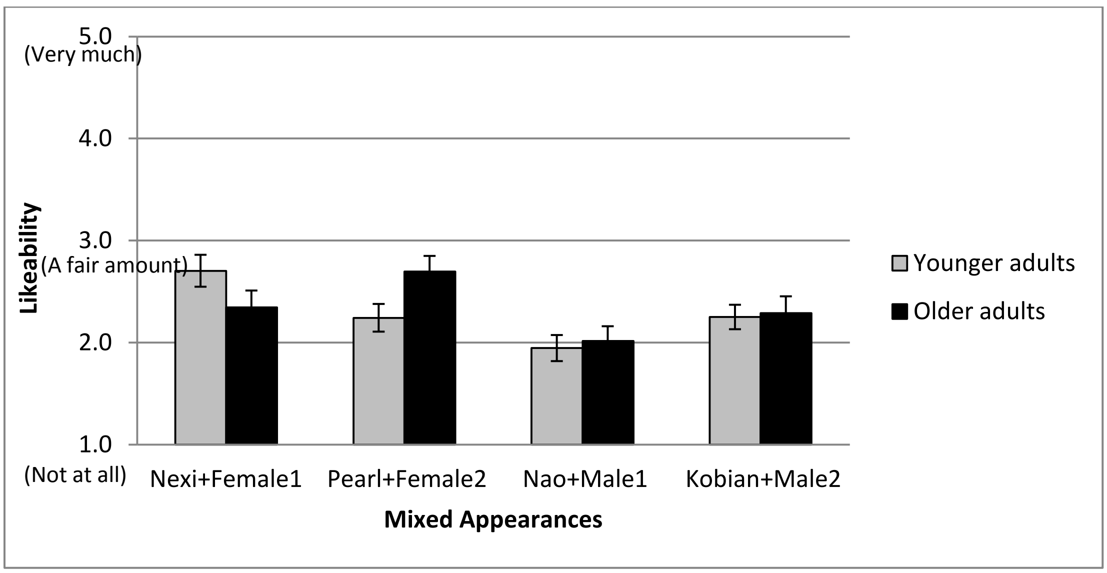


Figure 3.17. Mean likeability ratings for different mixed appearances. Error bars represent standard error of the mean.

Finally, an age X robot appearance (4) split plot ANOVA also yielded significant main effect of robot appearance (F (3, 186) = 27.23, $p < 0.001$, $\eta_p^2 = 0.30$) and significant age X robot appearance interaction (F (3, 186) = 10.21, $p < 0.001$, $\eta_p^2 = 0.14$). Post-hoc comparisons revealed that collapsed across age, Kobian was the least liked robot ($p <$ 0.001) and Nao was liked significantly more than Pearl ($p = 0.016$). However, when younger and older adults' data were analyzed separately, it was observed that younger adults liked Nao the most and Kobian the least (p's < 0.0167). Older adults also liked Kobian the least (only marginal significance reached when compared to Nao); however, they liked Nexi more than Nao (marginal significance; p = 0.03). See Figure 3.18 for a graphical comparison of mean likeability ratings for the four robots.

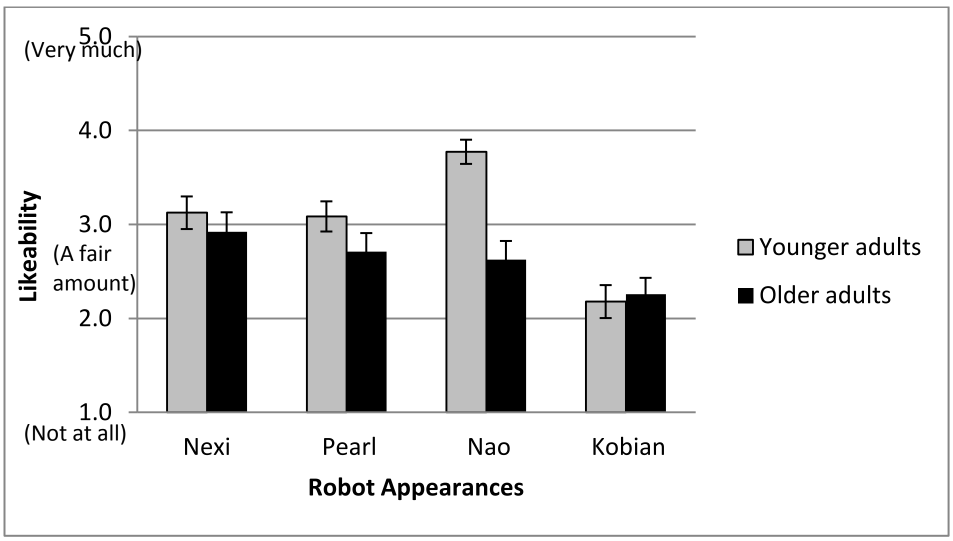


Figure 3.18. Mean likeability ratings for different robot appearances. Error bars represent standard error of the mean.

Of all the 12 appearances, Female2 was the only appearance for which older adults' mean likeability (M = 3.23, SD = 1.18) exceeded 3 (where 3 = a fair amount). Younger adults' mean likeability exceeded the mark of 3 for Female2 (M = 3.15, SD = 0.90) as well as for three of the four robot appearances (Nexi, Pearl, and Nao). Younger adults' mean likeability was highest for Nao (M = 3.77, SD = 0.73).

Summary of the Main Effects of Age, Humanness, and Task on Likeability

There were significant main effects of task and appearance on likeability. Robots were liked more for assistance with chores than with personal care and decision making. Additionally, mixed appearance was liked less than the human and robot appearances. Participants' likeability also varied within the human, mixed, and robot appearance categories. Of the human faces, Female2 was the most liked appearance. Of the mixed appearances, Nao+Male1 was least liked by younger and older adults both. However, the two age groups differed in their most liked mixed appearance – older adults liked Pearl+Female2 the most, whereas younger adults liked Nexi+Female1. Finally, in the robot appearance category, Kobian was the least liked robot for both the age groups. However, while Nao was the most liked robot for younger adults, older adults' likeability for Nexi was significantly higher than that of Nao.

<u>Overall Summary of the Effects of Age, Humanness, and Task on Likeability</u>

Humanness and task had an interaction effect on participants' likeability such that human appearance was least liked for assistance with a personal care task than with other tasks. Specifically for personal care task, likeability was higher for robot appearance than for human appearance. On the other hand, robot appearance was least liked for a decision-making task when compared to all other tasks. In fact decision-making task was the only task for which likeability for robot appearance was not significantly higher than that for mixed appearance. Averaged across the four tasks, mixed appearance was the least liked appearance (which explains the main effect of humanness).

A main effect of task was also observed for likeability suggesting that robot assistance would be liked more for chores than for personal care and decision-making. Finally, as was noted for PU and trust, likeability for human, mixed, or robot appearances also varied *within* every level of humanness.

Interpretation of the Rating Task Results

Although the age X humanness x task MANOVA on the linear combinations of perceived usefulness, trust, and likeability yielded significant interactions of age X humanness and age X task (besides other significant interaction and main effects), these two interactions were not significant for any of the DVs analyzed separately. This discrepancy results from the caveat associated with conducting univariate F-tests (ANOVAs) on multiple DVs – the univariate F-tests assume that there is no correlation among the DVs (Haase & Ellis, 1987). Perceived usefulness, trust, and likeability were, on the other hand, found to be highly correlated variables. MANOVA allows for the assessment of multiple correlated DVs together and therefore is the appropriate starting point of analysis in this case. Univariate analyses were conducted as follow-up tests on significant omnibus MANOVA results.

Univariate analyses of the effects of age, humanness, and task resulted in some common patterns across the three DVs. These patterns are summarized below:

1. Of all the tasks, robot appearance was evaluated least positively for decision-making. Decision-making was the only task where robot and mixed appearance evaluations were comparable (for both the age groups); for all other tasks, robot appearance was evaluated more positively than the mixed appearance. Moreover, of all the tasks, robot appearance was evaluated most positively for assistance with chores.
2. Of all the tasks, mixed appearance was evaluated more positively for chores. Younger adults (but not older adults) also evaluated this appearance more positively for decision-making compared to personal care and social tasks. Across both age groups, mixed appearance was evaluated less positively for assistance with personal-care and social tasks.
3. Of all the tasks, human appearance was *liked* the least for personal care. However, perceptions of usefulness and trust were not as low for human appearance in the personal care context. Compared to mixed appearance, human appearance was

evaluated more favorably for social task. Moreover, older adults, in particular, evaluated human appearance more favorably than mixed appearance for assistance with decision-making.

4. Robots (averaged across all appearances) were evaluated most positively for assistance with chores, particularly in comparison with personal care and decision-making tasks.
5. Overall, mixed appearance was perceived less positively compared to human and robot appearances.

Participants' perceptions were not uniform across all mixed appearances. Nao+Male1 was least positively evaluated by both the age groups. Nexi+Female1 was evaluated more positively by younger adults, and Pearl+Female2 by older adults particularly in comparison to Nao+Male1. Similarly, in the human appearance category, participants had most positive perceptions of Female2. In the robot appearance category, Kobian was evaluated least favorably. Younger and older adults differed in their most favorite robot appearance: while Nexi seemed to be a favorite of older adults (mean DV ratings around or above 3.0), younger adults made most favorable evaluations of Nao (mean DV ratings around or above 3.5). Pearl ranked second for older adults while Nexi was younger adults' second most favorite robot (see Appendix K for age by face mean ratings).

Of all the 12 faces, and across all DVs, older adults evaluated Female2 (human appearance) most positively, followed by Nexi (robot appearance). Younger adults formed most positive evaluations of Nao (robot appearance; highest mean trust and likeability) and Female2 (human appearance; highest mean PU).

Results Part 2: Qualitative Thematic Analysis

Overview of Thematic Analysis

To understand the reasons for participants' preferences for one appearance of robot over the others, the audio recordings of the 64 interviews were transcribed verbatim. The primary researcher developed a coding scheme based on the extant literature on robot appearance and social psychology (Table 3.3). The coding scheme was used to categorize the reasons participants gave for selecting a particular robot appearance. If participants' reasons did not fit into any of the categories of the coding scheme, the coding scheme was modified to be inclusive of the new response. The primary coder and a secondary coder coded the same two transcripts using MAXQDA text analysis software and were in 100% agreement. Thereafter, the remaining interviews were analyzed only by the primary coder.

Table 3.3

Primary Coding Scheme for the Analysis of Interview Data

Code	Definition
Humanness	Perceived humanness or machine-likeness of the robot
Gender	Actual or ascribed gender of the robot
Aesthetics/Design Features	Includes comments on specific facial features (e.g., eyes, nose, hair) and overall facial appearance (e.g., attractive versus ugly-looking)
Expressiveness/Personality	Includes comments on facial expressions (e.g., happy, sad, angry) and perceptions of personality (e. g, trustworthy, friendly, cute)
General Capability/Ability	Includes perceptions of physical (e.g., strong) and cognitive capabilities (e.g., smart, intelligent)
Other Reasons	Reasons that do not fit into any of the above codes

Most Preferred Appearance: Human, Mixed, and Robot

During the brief interview conducted at the end of the rating task, participants were asked to imagine having a robot helping them in their home with various tasks. With that imagination, they selected their most preferred human appearance, mixed appearance, and robot appearance and provided reasons for their preferences. Frequency distributions of participants' most preferred human, mixed, and robot appearances are shown in Figures 3.19, 3.20, and 3.21 respectively. Younger and older adults' preferences in the three categories matched the trends of means observed in the rating task.

Which Human Appearance was Preferred and Why?

A large majority of younger adults (84%) and older adults (78%) selected Female2 as the most preferred face for their robot in the human appearance category (Figure 3.19). The most striking reason that participants gave for this preference was the robot gender. Majority of participants mentioned that they would prefer assistance from a female robot in their homes than a male robot. Some older adults also perceived this face to be like that of a nurse or a caregiver (e.g., "*she looks like she's a nurse. Got her hair back and prepared to do the work.*"; "*Whereas she looks like a caregiver, someone that could be in with you, close to you.")*. On the other hand, some younger adults (both male and female) perceived a mother-like resemblance in the appearance (e.g., "*She reminds me of my mom. So that's kind of a deciding factor, it's familiar."; "especially if like the bathing thing, it seems like a motherly aspect, so it'd be more comfortable than like another guy, I guess, in the room."*).

The aesthetics of the face (e.g., the eyes, hairstyle) also influenced this selection. Although all faces had neutral expressions, some participants perceived a hint of smile on Female2's face which also affected their preference. Some participants who preferred a male face attributed more intelligence to a male-looking robot for investment-related tasks (decision-making). Other reasons for preferring a male face were perceptions of more strength in a male-robot, and in case of a few male participants, the comfort expected from a same-gender robot (e.g., a male older adult mentioned, " *if they were going to bathe and be with me 24/7 I would want, probably a male. And if it wasn't a robot, if it was a nurse, I'd probably pick a male nurse over a female nurse*.")

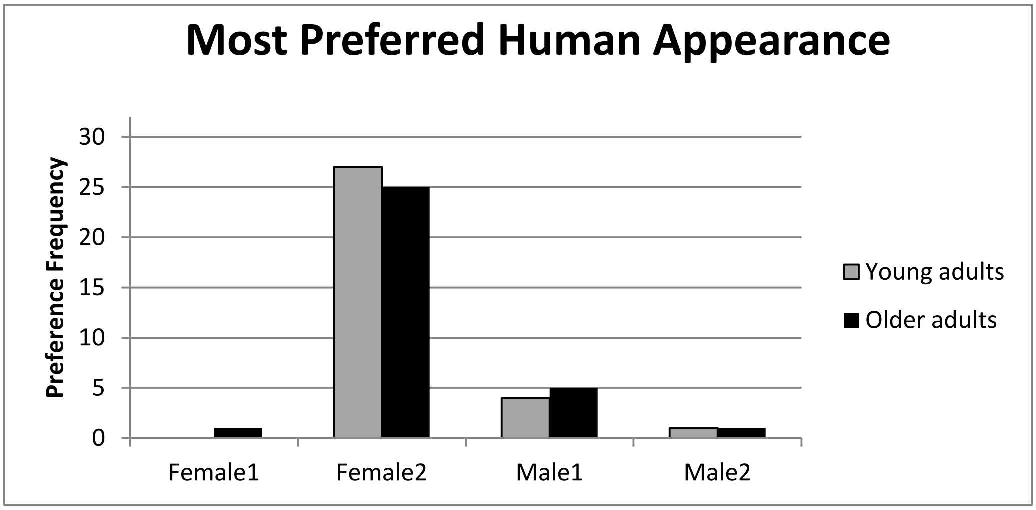


Figure 3.19. Younger and older adults' most preferred human appearance.

Which Mixed Appearance was Preferred and Why?

Younger and older adults differed in their most preferred mixed appearance (Figure 3.20). While Nexi+Female1 was most popular younger adults, older adults showed higher preference for Pearl+Female2. Younger adults preferred Nexi+Female1 primarily because it was more human-looking than the other mixed-appearances. The younger age group also ascribed more intelligence to this appearance which further led to their preference of this face.

In the older age group, 50% participants preferred Pearl+Female2 and 37% preferred Nexi+Female1. Pearl+Female2 was preferred primarily for the aesthetics and perceived personality (e.g., pleasant, companion-like) of the appearance. About half of the older adults who preferred Pearl+Female2 made specific comments about its eyes such as, "*The eyes are telling me this robot can be trusted…there's just something about the eyes that just make, I mean when I communicate I look at a person's eyes…"*.

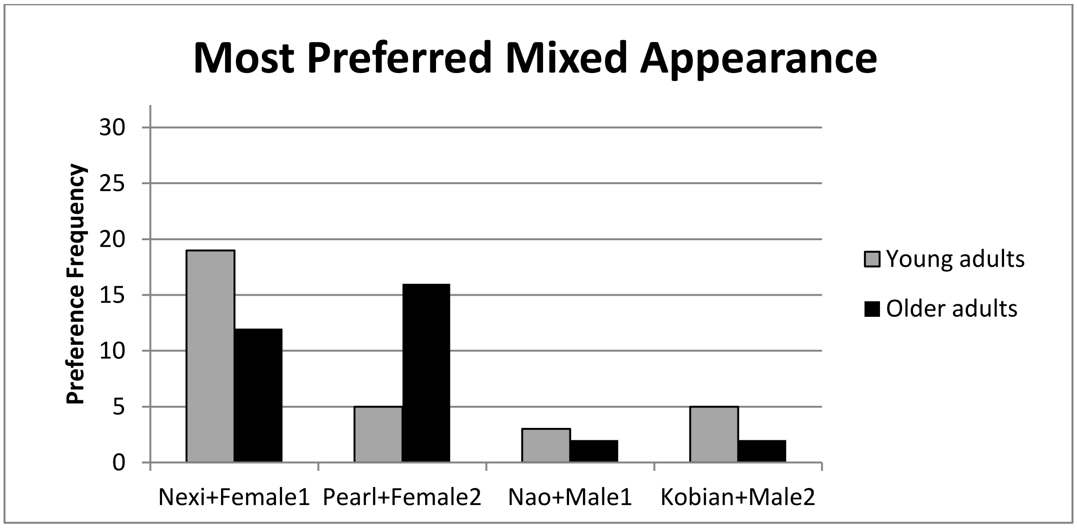


Figure 3.20. Younger and older adults' most preferred mixed appearance.

<u>Which Robot Appearance was Preferred and Why?</u>

More than half of the younger adults preferred Nao over the other robot appearances (Figure 3.21), 7 of whom considered its appearance to be "cute". Another primary reason given for the preference of Nao was its neutral expression which seemed to fit with how many younger adults imagine their robot to look like (e.g., *"[Nao] just looks like a normal, when I picture a robot, that's what I picture the white face with the eyes"; [Nao] looks like very constant…like they don't have as much emotion."; [Nao] seems like it's like trying to be more robot-y…I like that [it] isn't trying to be a human.*) Not having too many facial features was considered a plus, as can be observed in this comment, "*[Nexi] is kind of reminiscent of a little brother or a little kid, and it has to do these chores, and I'd kind of feel bad. [Nao] doesn't really have the ability to move its mouth or raise its eyebrows, so I'd never know… which may not be good. Yeah. That feedback isn't there, which is kind of a plus."*

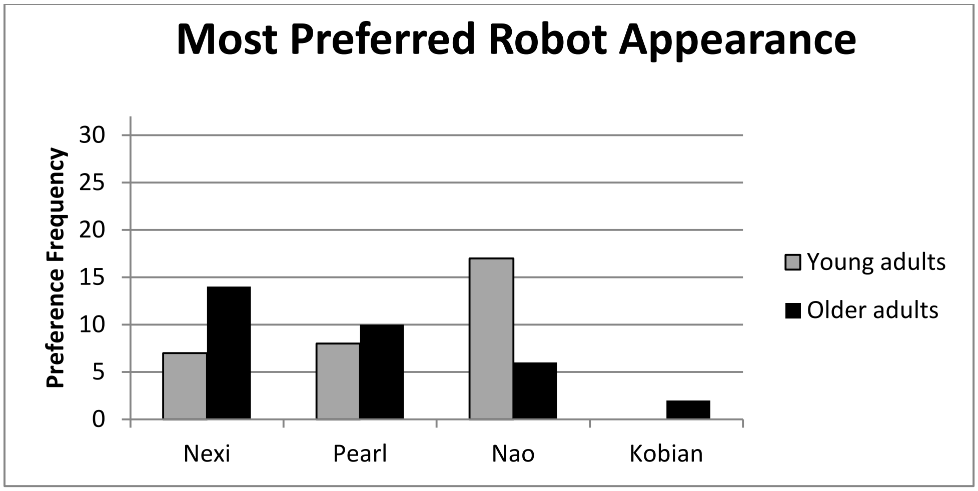


Figure 3.21. Younger and older adults' most preferred robot appearance.

Nao was not as popular among the older adults as it was among the younger adults. Forty-four percent of the older adults preferred Nexi over the other robot appearances. A common reason was the expressiveness of the appearance resulting from the combination of the facial features, as can be illustrated through this comment, "*I like [Nexi's] big eyes. Uhh, I like his round face instead of that square face. I like the fact that he has a real mouth instead of a line. This looks like a brain centre that he could be think about. His eyebrows look like they could move and make an expression on his face.*"

Global Preference: Human versus Mixed versus Robot

Of their most preferred human, mixed, and robot appearances, participants selected their most preferred appearance overall and answered why they preferred that appearance over others. The frequency distribution of preferences is illustrated in Figure 3.22.

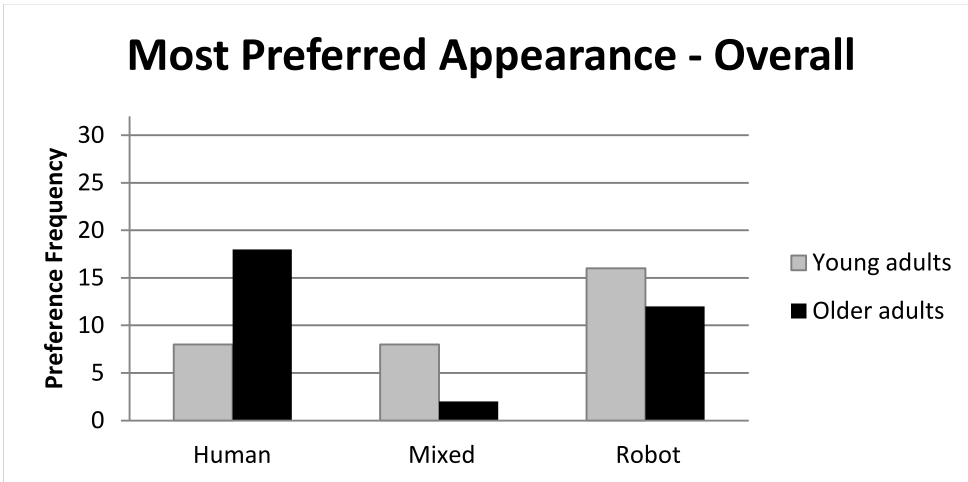


Figure 3.22. Participants' selection of their most preferred face for their robot.

A chi-square analysis to assess whether the distribution of the most preferred face across human, mixed, and robot appearance categories depended on the age was significant (χ^2 (2, N = 64) = 8.02, $p < 0.05$). Of all the participants who preferred human appearance, about 70% were older adults. Chi-square analyses were also conducted to assess if human, mixed and robot appearances would be preferred by an equal number of people. This analysis was conducted separately for older and younger adults. Older adults' preferences were on the extremes, that is, they were in the favor of a completely human-looking appearance (56%) or a robotic appearance for their robot (37%), but would not prefer a mixed appearance (χ^2 (2, N = 32) = 12.25, $p < 0.05$). However, younger adults' preferences were more varied; half of them preferred robotic appearance, a fourth preferred mixed appearance and another fourth selected human appearance (χ^2 (2, N = 32) = 4.00; $p > 0.05$).

Why People Prefer Highly Humanlike Appearances?

Fifty-six percent of the older adults and 25% of the younger adults selected a human face as their most preferred appearance for their home robot. Their major reason for their selection was the human-likeness of the face. Many participants elaborated this further and mentioned that a human-like face would be more familiar and "relatable" than other robotic appearances, for example, "*I guess that whole idea of having a robot kind of freaks me out a little bit. Um, so yeah. I like that it looks like a human. I feel like I could connect better with it.*" Moreover, people also considered a human-like appearance more apt for fulfilling companionship needs, as is exemplified in this remark from an older adult, "*it's not only capable of doing chores and functionally but it has within it the capability of being a companion. So your companions look more like you than or resemble you, something that's familiar and that one does it."*

Some participants perceived a human-like robot to be more capable in general than other robots, for instance, "*...she's more able to perform the duties she's supposed to, I guess."* In fact, a few participants considered human-looking robots to be "the most developed" kind of robots. In some cases, preferences were influenced by the robot gender as well: "*Uh it's a lady, um and that would be just a good companion or somebody for me to talk to, work with…".* On the contrary, another participant who favored a male human appearance said, "*And I know you said each one can do everything, but this one kind of is uh, probably preconceived idea that the man-looking robot would be able to do everything."* The perceived personality or expressiveness of the human face was also pointed out as reasons for preference as noted in descriptors such as "smart", "caring", "non-intimidating", and "[has] a little smile".

<u>Why People Prefer Less Humanlike Appearances?</u>

About 38% older adults and 50% younger adults selected a robotic face as their most preferred appearance for their robot. Most of these participants did not want their robot to resemble a human because it would be difficult for them to distinguish such a robot from a human: "*Well, when I think of a robot, I think of him not being people. Umm these look very realistic and I might confuse a person with a robot. This [the selected picture] is definitely a robot and this I would be in command of."* People also preferred robotic appearance due to the perceived personality or expressiveness of the face (e.g., they found the appearance to be "cute", "friendly", "child-like", and/or "trustworthy".)

Robotic appearance was also favored by participants who tended to ascribe negative human traits or intentionality onto human-looking robots, for instance, a person reasoned why she would not like a human-looking robot, "*She just looks like she could tell me a lie. Just be like Yes ma'am, yes ma'am, and in the back of her mind she's like 'I can't wait to get out of here.'*" Such participants perceived robotic appearance robots to be devoid of such flaws: "*It seems like it will do exactly what it's supposed to do.*"

Only about 6% older adults in our sample preferred a mixed appearance over human and robotic. However, a considerable proportion of younger adults (25%) had most preference for a mixed appearance. These participants explained that a mixed appearance was better than the extremes because it employed the benefits of human and mechanical appearance. A quotation from a young adult exemplifies this reasoning, "*…because although it's human enough to be familiar, it's, like, clearly not human so…I still perceive it as like a robot, but it doesn't make me as uncomfortable as a human face would on a robot.*" A similar justification was noted in another participant's comment: "*It's not quite as I guess invasive as having another person living with you, but it's not as unrealistic as having like a robot from a horror movie or something living with you. It's a good blend of both".*

Many participants, particularly older adults, who did not favor mixed appearance spoke against its aesthetics or design features and compared it to "alien-like" appearance, for example, "*that space thing on the head looks like something from outer space…*"; "*Looks like a space man or space woman…"; "Might help if he wasn't bald. Looks sort of alien without having any hair…"* Interestingly, while many younger adults liked the human-machine blend in the mixed appearance, some older adults used the same reason to not like this appearance, for instance, "*But this one, kind of gets the worst of both. It's not as pleasant as this one and not as familiar as this one.*"

Task-Specific Preferences for Appearance

After assessing global preference for appearance, task-specific preferences were examined. Participants' preferences for robot appearance varied across different tasks. For example, for chores, there were less distinct preferences for any kind of appearance (a considerable proportion opted for no preference). However, when asked to think specifically about a social task 60% of older adults and 50% of younger adults preferred human face for their robot. Perceptions of human-like "sociability" could have led to this preference. For example, an older adult reasoned, "*because she's ah, well it's, she looks more capable of being sociable than these two. She looks more human-like."* Similarly, a younger adult, who preferred a human appearance for this task, commented, "*particularly with a social task, you wanna be dealing with a human. Or at least, make it seem like you're dealing with a human more than a robot.*"

Of the four tasks, decision-making task is the one for which mixed appearance was evaluated more favorably, particularly by younger adults (see Figure 3.23). Perceptions of "intelligence", "wisdom" and "smartness" often influenced this selection. Some example statements are: "*it's just the whole intelligent look…"*, "*because he seems more wiser with the glasses*", and "*that one looks more smart"*.

Preferences seemed divided between human and robotic appearance for a personal care task such as bathing. Those who preferred a human appearance associated more human-like care and capabilities with such a robot: "*I'm just more comfortable with this robot that looks more like a nurse or a nursing assistant. It looks like a humanoid that you could trust and I'm giving them the benefit of knowing how to aid and hold you as immerse into the water bathing or that sort of stuff*". On the contrary, many others did not want a human-looking robot to perform a task so personal in nature, as was reflected in the comment, "*sometimes personal care can get pretty involved, and I'd much rather have an impersonal looking creature caring for my personal needs.*"

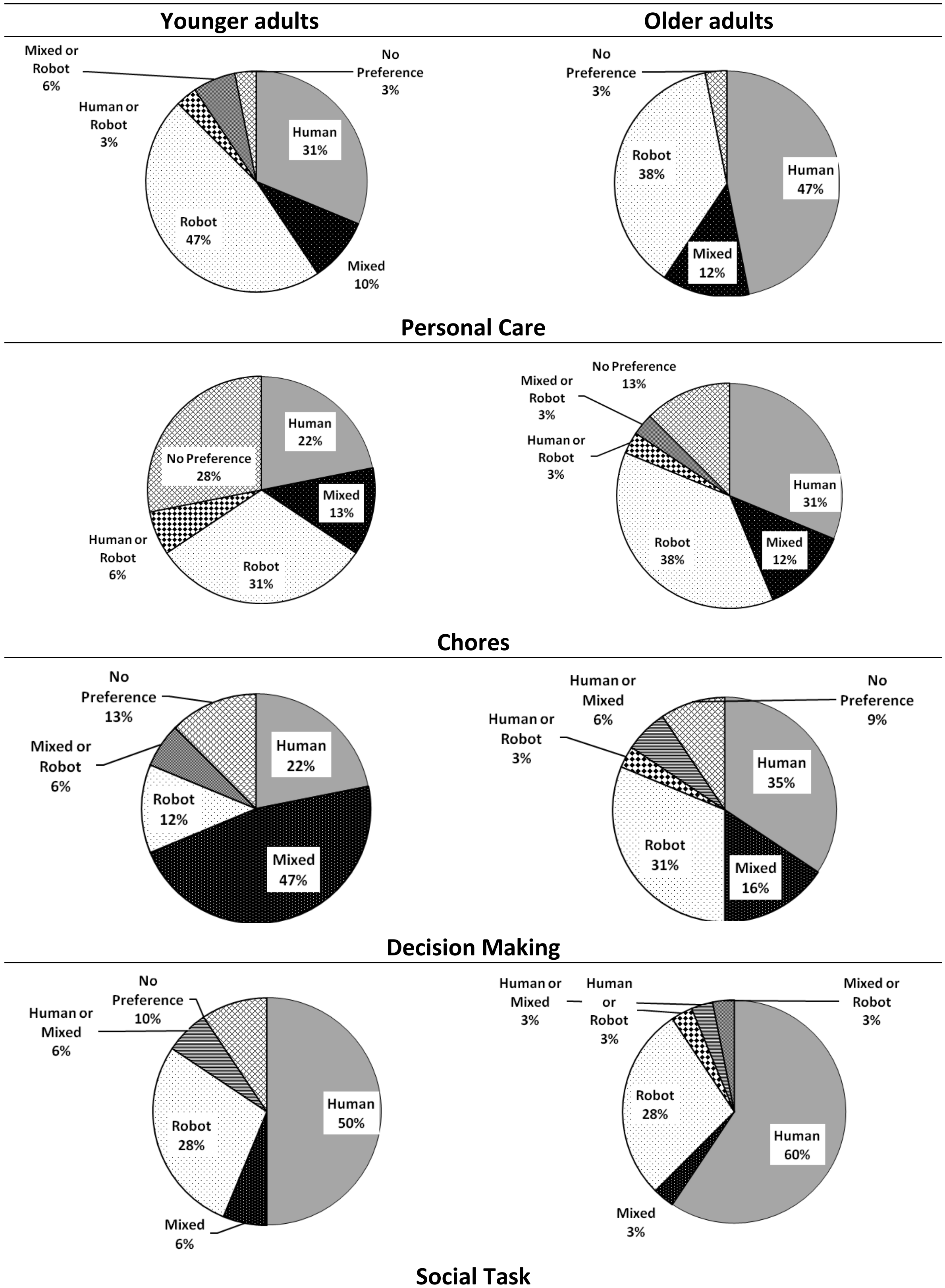


Figure 3.23. Distribution of appearance preferences across tasks.

Results Part 3: Analysis of Questionnaires

Robot Opinions Questionnaire

Participants' attitudinal acceptance of robots was assessed through a robot opinions questionnaire (Smarr et al., 2012). Younger (*M* = 5.31, *SD* = 0.95) and older adults' (*M* = 5.32, *SD* = 1.08) mean responses on the questionnaire were comparable. Both groups had generally positive opinions about robots (7 point scale; 1 = extremely unlikely, 4 = neither, 7 = extremely likely).

Robot Familiarity and Use Questionnaire

Participants' familiarity with and use of 13 different kinds of robots was measured on a 5-point scale (0 = not sure what it is; 4= have used or operated this frequently). Younger adults' reported familiarity with robots was significantly higher than older adults (t (62) = 2.73, $p < 0.05$). Both the age groups reported least familiarity for remote presence robots (e.g., Texai, Anybot). Older adults reported most familiarity for manufacturing robots (e.g., robotic arm in factory) whereas younger adults' reported familiarity was highest for entertainment/toy robots (e.g., Aibo, Furby). For most robots, the average familiarity ratings implied that participants either had no idea of the robot in question or had only heard of or seen it, but did not have experience using it (Appendix L).

Robot Facial Appearance Questionnaire

The facial appearance questionnaire, consisting of 15 items, was designed for this study to assess the facial features and facial characteristics that people want their robot to have. Mann-Whitney *U* test was performed on each of the fifteen items to compare younger and older adults' responses. Significant differences (tested at p < 0. 05) were observed on 5 items: I would want my robot to have lips; I would want my robot to have hair on its head; I would want my robot to have soft skin; I would want my robot to have a round face, and I would want my robot to look exactly like a human. For these five items, one-sample Wilcoxon signed rank test was also performed (separately on the two age groups) to assess if younger and older adults differed from 3 (i.e., the "neither agree nor disagree" point) and in which directions.

Compared to younger adults, older adults were more likely to want their robot to have lips (Z = -3.35, p = 0.001). When compared to the neutral point (= 3), older adults were found to be significantly above (Z = - 4.28, p < 0.001), implying that this age group wanted their robot to have lips. Younger adults' response was not significantly different from the neutral point (Z = - 1.89, p = 0.06).

In comparison to younger adults, older adults were also more likely to want their robot to have hair on the head (Z = -3.41, p = 0.001). The 1-sample Wilcoxon test was non-significant for older adults but was in the positive (agreement) direction (Z = - 1.93, p = 0.05). The 1-sample Wilcoxon test was significant for younger adults (Z = -2.51, p = 0.02) in the negative direction, implying that younger adults *did not* want their robot to have hair on its head.

On the item "I would want my robot to have soft skin", older adults were again more likely to agree than younger adults (Z = -2.03, p = 0.04). The 1-sample Wilcoxon tests yielded significant result for older adults in the agreement direction (Z = -2.48, p = 0.01) whereas non-significant result was obtained for younger adults (Z = - 0.29, p = 0.77). Thus older adults wanted their robot to have soft skin whereas younger adults neither agreed nor disagreed.

Younger adults were more in agreement to having a round faced robot than were older adults (Z = -2.40, p = 0.02). The 1-sample Wilcoxon test showed that both younger (Z = -4.21, p < 0.001) and older adults (Z = -2.60, p = 0.01) wanted their robot to have a round face.

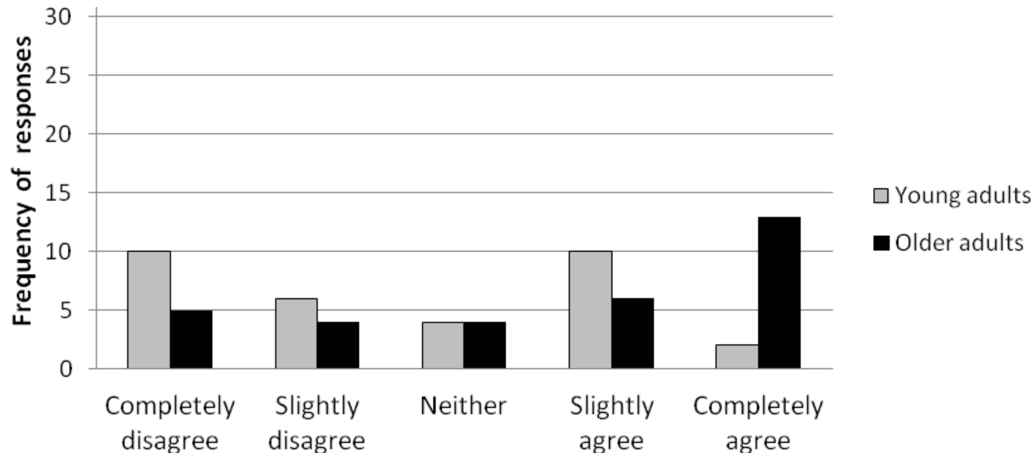


Figure 3.24. Frequency distribution of responses on the item “I would want my robot to look exactly like a human.”

Younger and older adults’ responses were also significantly different on the item “I would want my robot to look exactly like a human” (Z = -2.61, p = 0.01; Figure 3.24). Whereas older adults agreed that they would want a robot that looked exactly like a human (1-Sample Wilcoxon Z = -1.97, p < 0.05), younger adults tended to disagree although their median was not significantly different from the neutral point (Z = 1.72, p = 0.09).

CHAPTER 4

DISCUSSION

In human-human interaction, facial appearances influence formation of initial impressions. One of the primary goals of this study was to investigate if initial perceptions formed towards robots would also be influenced by the humanness of the robot's face, particularly when the robot is providing assistance with tasks that are traditionally carried out by humans. Moreover, although robots have the potential to help both younger and older adults, there is limited knowledge on how the two age groups' perceptions of robot humanness compare with each other. Therefore, an additional goal was to examine if younger and older adults differed in their perceptions.

At a general level, a mixed human-robot facial appearance was evaluated less positively than a highly human-looking or a highly robot-looking appearance. This trend was observed in the rating task across the measures of perceived usefulness, likeability, and trust for both younger and older adults. This finding seems aligned with the uncanny valley theory (Mori, 1970), implying that a robot face that partially imitates a human appearance evokes less positive perceptions than a more mechanical or a completely human-like robot face. However, one of the caveats of the earlier research on uncanny valley theory was the ill-defined context in which robot appearances were evaluated (e.g., MacDorman, 2006; MacDorman & Ishiguro 2006). This caveat was addressed in the current study by asking participants to imagine interacting with the robot in specific task contexts.

When the task was taken into account, the trends in perceptions were more complex and deviations from the uncanny valley pattern were observed. For example, robot (mechanical) appearance was evaluated more positively than the mixed appearance for chores, social, and personal care tasks. However, for decision-making task, mean ratings for robot appearance were comparable to those for the mixed appearance. Additionally, age-related differences in perceptions were also noted. The younger adults (but not the older adults) evaluated the mixed appearance more positively for assistance with decision-making than with personal-care and social tasks. Thus, this study evaluated perceptions of a broader range of users and found differential perceptions across age.

Prior research on robot appearance that did take robot task into consideration did not assess the underlying reasons for the preference of one appearance over the other (e.g., Goetz et al., 2003) The multi-method approach used in the current study identified not only the patterns of perceptions across different appearances but also the possible reasons that influence the formation of such perceptions. For example, in the rating task participants' perceptions for the robot (mechanical) appearance were found to be least favorable for the decision-making task. The interview data revealed that participants varied their evaluation criteria for robot appearance across different tasks.

For the decision-making task, the appearance that evoked perceptions of intelligence, smartness, or wisdom was preferred for assistance. Perceptions of "cuteness" or "friendliness", which were frequently mentioned as reasons for a general preference of the mechanical appearance, were not considered important when evaluating assistance for a cognitively demanding task such as decision-making. For a considerable proportion of the younger adults, a mixed appearance with an appropriate blend of human-mechanical appearance met the criterion of intelligence, and was preferred over the other appearances for decision-making. However, older adults considered the mixed appearances to be less familiar or alien-like, and were therefore, more in favor of the human appearance for this task.

The results of this study have implications both for advancing theoretical understanding of robot perceptions and for creating and applying guidelines for the design of robots. These are discussed separately in the next sections.

Theoretical Implications

As measured via the robot opinions questionnaire, both younger and older adults had generally positive opinions about using robots. However, compared to the younger adults, the older adults had less familiarity and experience with robots (assessed through the robot familiarity and use questionnaire). Therefore, the older adults' perceptions about a robot's appearance were more likely to be shaped by their expectations than by past experiences with a robot. Older adults' higher preference for a human appearance could be an outcome of such inexperience.

A primary reason for why human-looking robots might be favored over mechanical appearance is *familiarity* with the human appearance, particularly for performing tasks in the home that typically have been performed by humans (Blow et al., 2006). However, such participants might assume the robot to be a perfect copy of a human, triggering the same nuance of familiarity as is evoked by another human. From the perspective of the uncanny valley theory (Mori, 1970) this would happen when the second peak of the graph is reached (see Figure 1.1). The primary proposition of the uncanny valley theory is also based on the notion of familiarity. The argument is that as a robot is designed to appear more human-like, the familiarity with it reduces because the appearance seems to be a faulty replica of a human. However, if the human-like appearance can be perfected to reach a point where it is indistinguishable from human appearance, positive perceptions (due to high familiarity) will be evoked.

Although the majority of the older adults preferred a human appearance, another considerable proportion of them (37%) leant toward the least human-looking appearances for their robot. Many participants, including younger and older adults, raised concerns about not being able to differentiate a robot from a human. The closer the robot's face resembles to a human, the more likely it is to be anthropomorphized through overgeneralization effects (Epley, Waytz, & Cacioppo, 2007; Zebrowitz and Montepare, 2008). Thus it would be difficult for the human users to inhibit attributions of human strengths and weaknesses onto a human-looking robot. Moreover, people would already have expectations about the behavior of the robot. This would apply even to people who prefer a human appearance for their robot. For example, some participants assumed that a human looking robot would be more capable than other robots, and therefore favored the appearance. On the contrary some participants attributed human flaws (such as disobedience and betrayal) onto human-looking robots and therefore were more inclined toward mechanical appearance. Therefore, familiarity-based overgeneralizations could be both beneficial and problematic for the acceptance of human-looking robots. Thus, theoretical models of robot perception need to be inclusive of the positive and negative effects of familiarity and expectations that emerge from a human appearance.

Robot appearance research has focused on identifying general patterns in perceptions and preferences. These patterns predict a trend of behavior for most people but ignore or undervalue inter-individual differences. For instance, in the current study, more positive evaluation of the human and the robot appearance over the mixed appearance offered support for the uncanny valley phenomenon at a nomothetic level. However, a few participants (25% of the younger adult sample) chose a mixed appearance as their *most preferred* appearance during the interview. Such participants might have a different trend of perceptions across varying levels of robot human-likeness. This means that the uncanny theory even if validated at a nomothetic level, might not hold true at an ideographic level.

One of the sources of the inter-individual differences can be the participants' age cohort. The present-day older and younger age groups differ not only in their direct experiences with robots but also in their exposures to robot-specific scientific fiction (e.g., novels, movies, and TV series). Such differences in experience could lead to different expectations toward robot appearance. Thus, individual differences should also be incorporated into a model of robot perceptions by systematically considering a wider range of potential users.

Previous works on robot appearance have also underestimated *inter-individual* differences in appearance perceptions by not taking the context of human-robot interaction into account (e.g., Bartneck, Kanda, Ishiguro, & Hagita, 2009). The present study found that task context affected individuals' perceptions of robot humanness. Thus, individuals calibrated their appearance preferences based on the attributes of the task. Therefore, perceptions and preferences gauged in isolation with task contexts are less informative of the participants' evaluation criteria and judgment processes.

In addition, in the extant literature on humanoid appearance much emphasis has been on comparing the effects of different levels of humanness (e.g., MacDorman, 2006; MacDorman & Ishiguro 2006). The assumptions of such research undermine the variability *within* a particular level of humanness. It is not merely the degree of humanness but also specific characteristics such as gender, expressiveness, aesthetics, and perceived capability or intelligence that influence people's perceptions. Moreover, tasks are also stereotyped by gender, which can further influence perceptions of male versus female looking robots (Heilman, Wallen, Fuchs, & Tamkins, 2004). People in the present study were more likely to prefer a female human-looking robot for general assistance in their homes. Assistance with chores, personal care, and social companionship are tasks stereotypically associated with females. However, participants' comments indicated that for more cognitively demanding tasks (e.g., decision-making in managing finances, other male gender-typed tasks), preference for assistance might shift toward a male-looking robot. Additionally, some male participants leant toward a male-robot for assistance with personal care tasks.

Robotic (mechanical) appearance was preferred overall when it was perceived as "cute", "friendly", "trustworthy" and/or easy to command. However, all mechanical appearances were not perceived equally favorably. Similarly, even though the mixed appearance was less positively evaluated, there were differences in perceptions within that category. Therefore, specific characteristics of any robot appearance (gender, aesthetics/features, expressiveness, etc.) may interact with the robot's humanness to affect robot acceptance.

Applied Implications

From the applied perspective, it is important to consider younger and older adults as heterogeneous groups. Thus, although, on average, older adults may show a higher inclination toward human-like appearance of robots, differences in preferences *within* the two age groups should not be overlooked. A considerable proportion of older adults also preferred robotic (or mechanical appearance). Younger adults' preferences were even more diverse – half of them preferred robotic appearance, a fourth preferred human appearance and the remaining fourth preferred mixed appearance. Thus, in general, there is wide variability in people's attitudes toward mechanical and humanoid appearances. Overall, robot designers should consider designing robot appearances keeping the following four categories of users in mind (Figure 4.1):

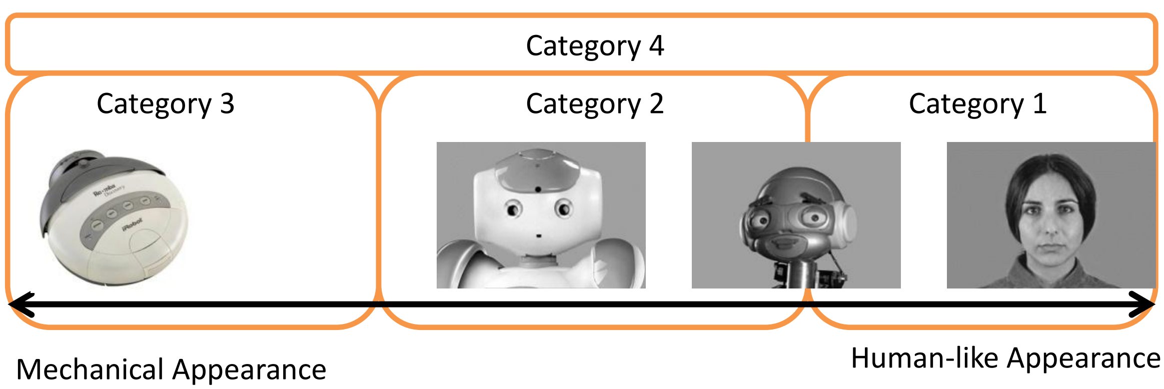


Figure 4.1. Depiction of four categories of users based on preferences for robot appearances, from mechanical to human-like; Category 4 overlaps with all appearances assuming same functionality across appearance.

- **Category 1**: Those who want their robot to look like a human or at least have some human characteristics.

- **Category 2**: Those who do not necessarily want their robot to look like a human; but would find some humanoid characteristics acceptable.
- **Category 3**: Those who do not want their robot to look like a human *at all*.
- **Category 4**: Those to whom appearance does not matter; functionality does.

The primary reasons associated with a preference for human-like appearance (Category 1) are familiarity, ease of interaction, perceptions of higher capability, and expectations of companionship from the robot. As was noted in the present study, people in this category also consider human-looking robots as the most "developed" kind of robots and expect the robot to be more than a mechanical tool.

In contrast, individuals who do not want their robot to be human-like (Categories 2 and 3) may be concerned about confusing a robot with a human being. Some of them even tend to attribute human flaws (such as disobedience and betrayal) to human-looking robots. Category 2 comprises older adults who do not want the robot to be highly human-looking but would not mind some humanoid features particularly if the robot appears "cute", "friendly", "trustworthy" and/or "easy to command".

People in Categories 3 and 4 view a robot only as a functional tool. However, Category 3 people want their robot to look like a machine whereas Category 4 would not be concerned about the looks of the robot; if humanoid robot A can perform a task as well as the non-humanoid robot B, Category 4 individuals would be equally receptive of A and B.

Another important consideration in robot face design is about the composites of the face that make it desirable versus non-desirable for a particular user-group. Nexi+Female1 (see figure 2.1) was the most preferred mixed appearance for many younger adults and was also selected as the most preferred appearance overall. Noticeably, this stimulus was bald. Based on the response on robot facial appearance questionnaire, we found that younger adults *did not* want their robot to have hair on its head and their responses were significantly different from the older adults in the opposite direction. Moreover, the older adults wanted their robots to have lips and soft skin, whereas these features were not deemed as desirable to younger adults. Both younger and older adults wanted their robot to have a round face and older adults desired it even more than younger adults. Furthermore, older adults were more likely to want their robot to be completely human-looking whereas younger adults tended to be neutral in their response to this question.

People's evaluation criteria for an appearance (e.g., cute versus smart) varied across task. Thus, if the robot is designed to specifically assist the user with critical decisions, the robot should not be given a funny demeanor, but an intelligent look. For assistance with a personal care task (e.g., bathing), users can be divided into two broad groups: those who would prefer the human appearance and those who would prefer a mechanical/robotic appearance. Human appearance is preferred for assistance with personal care because it evokes in the user perceptions of "nurse-like" capabilities. Thus, people may be more comfortable in taking assistance from the robot. However, the human appearance might also be considered invasive for a highly personal task which is why a mechanical appearance would be preferred by the other user group. For a social task (chatting, providing companionship, helping learn a new skill, etc.) human-like sociability would be an important attribute of the robot appearance.

Methodological and Measurement Considerations

This study was designed to systematically investigate the effect of robot humanness on peoples' perceptions. One of the strengths of the current design was the use of multiple pictures for every level of humanness as it provided multiple data points at each level of robot humanness. Thus, we could assess if participants were uniform in their perceptions within every level of appearance, and we found that to be untrue. For instance participants had highly positive perceptions toward one of the human faces but not toward the others. This helps us understand that it is not only the humanness of the face but also the specific characteristics of the human appearance that impact perceptions.

Another strong aspect of the study was the assessment of robot appearance perceptions in the context of tasks. Robots can be designed to assist with various activities of daily living; however depending on the activity, the nature of robot's assistance would differ in terms of physical interaction with the user (e.g., proximal interactions for personal care versus distal for chores), cognitive support needed (high for decision-making versus low for chores), and social interactions (high for social task versus low for chores). Thus, people's perceptions of robot humanness would also likely be dependent on the nature of the assistance needed from the robot.

The use of multiple dependent measures was also an important extension of earlier HRI studies. Evaluating perceptions across different measures provided a clearer relation among the constructs that constitute perceptions and predict attitudinal acceptance of robots. Although the three measures of positive evaluations were highly positively correlated, perceived usefulness and trust were almost perfectly correlated for older adults. This implies that the robots that are considered more useful by older adults will also be trusted more and vice-versa. An alternative possibility is that people do not differentiate those terms and thus, they are not sensitive enough to differentiate attitudes on those dimensions.

Likeability was a measure of the affective component of participants' attitudes whereas usefulness was a measure of the cognitive component. Trust measures both affective and cognitive components. The uncanny valley theory was supported most by the patterns of the likeability ratings across the three levels of appearances and least by usefulness ratings, as was evidenced by the effect sizes. Thus, individuals' *affective* reactions toward robots seemed more variable across the three levels of appearances compared to their *cognitive* assessments (i.e., their perceptions of usefulness).

Despite the systematic design, the study was not without limitations. The biggest caveat was the lack of robot interactivity due to the use of still pictures. Moreover, although vignettes were provided for every task situation, participants were restricted to their imaginations to evaluate the nature of robot assistance with different tasks. Additionally, limitations were noted in using “anxiety” as a dependent measure. As was expected, anxiety was negatively correlated with perceived usefulness, trust, and likeability. However, less confidence can be put into the strength of the correlations due to the issues associated with the anxiety data (see Appendix J). Future studies should use such constructs to measure negative attitudes toward robots that are more easily and unambiguously understood by participants (e.g., fear, discomfort). Moreover, using bipolar Likert-type items, with a neutral-point in the middle, would allow for the measurement of positive to negative valence on the same dimension (e.g., comfort-discomfort, like-dislike, trust-distrust, useful-useless). In addition, it would minimize confusions regarding the direction of the scale.

The findings of this study should be carefully generalized. The older age group sample in the study was represented by relatively healthy and considerably educated older adults living in the Atlanta metropolitan area of the United States. Older adults with different backgrounds from the present sample might have different preferences and attitudes toward robots. Similarly, the younger age group was represented by the undergraduate students of Georgia Institute of Technology. The interactive effect of educational background and age-cohort experiences on attitudes toward robots was not explored in the current study. Therefore, the degree to which less educated younger adults would resemble Georgia Tech younger adults in their perceptions of and preferences toward robot appearance remains unclear.

Conclusion and Future Directions

People's perceptions of robot faces clearly vary as a function of robot humanness. In general, people perceived a mixed human-robot appearance less favorably compared to highly human and more robotic appearance. Additionally the nature of task also influenced people's overall perceptions of robots. Robots are most positively evaluated for assistance with chores and less positively for personal care and decision-making. This finding is consistent with previous studies (e.g., Smarr et al., 2012). Moreover, task and robot humanness have an interactive effect on people's likeability, trust, and perceived usefulness toward robots.

There are age-related differences in *preferences* of robot humanness. Older adults showed a higher inclination toward human-looking appearance of robots whereas younger adults' preferences were more distributed across the levels of humanness. An appearance with mixed human-robot features may be more likely to be rejected by older adults than by younger adults, and the difference would be most striking for a decision-making task.

Besides the humanness of the robot face, perceptions of robot appearances are also influenced by other factors such as robot gender, facial features/aesthetics, expressiveness, perceived personality, and perceived capability. Future studies should measure the relative weight of these different factors in the formation of perceptions, both at a global level and at a task-specific level. Moreover, an understanding of the aesthetics that evoke perceptions of intelligence, sociability, and/or humanness could further improve the design of robots.

APPENDIX A

COMMON MEASURES OF ROBOT PERCEPTION

Table A.1

Different Measures of Perceptions Used in Studies Investigating Uncanny Valley Theory

Name of the construct	**Reference of studies**
Affect evoked (e.g., fear, anxiety)	Nomura, Kanda, Suzuki, and Kato (2004)
Attractiveness	Chen, Russel, Nakayama, and Livingstone (2010); Hegel, Lohse, and Wrede (2009)
Familiarity	McDorman (2006); MacDorman and Ishiguro (2006)
Likeability	Bartneck, Kanda, Ishiguro and Hagita (2009); Groom et al. (2009); Hegel, Lohse, and Wrede (2009); Mathur and Reichling (2009)
Perceived Eeriness	MacDorman (2006); MacDorman & Ishiguro (2006)

APPENDIX B

DEMOGRAPHICS AND HEALTH QUESTIONNAIRE

Please answer the following questions. All of your answers will be treated confidentially. Any published document regarding these answers will not identify individuals with their answers. **If there is a question you do not wish to answer, please just leave it blank and go on to the next question.** Thank you in advance for your help.

Demographics Questionnaire

Gender: Male ☐$_1$ **Female** ☐$_2$ **Age: _______**

1. What is your highest level of education?

☐$_1$ No formal education
☐$_2$ Less than high school graduate
☐$_3$ High school graduate/GED
☐$_4$ Vocational training
☐$_5$ Some or in-progress college/Associate’s degree
☐$_6$ Bachelor’s degree (BA, BS)
☐$_7$ Master's degree (or other post-graduate training)
☐$_8$ Doctoral degree (PhD, MD, EdD, DDS, JD, etc.)

2. Current marital status (check one)

☐$_1$ Single
☐$_2$ Married
☐$_3$ Separated
☐$_4$ Divorced
☐$_5$ Widowed
☐$_6$ Other (please specify) __________________

3. Do you consider yourself Hispanic or Latino?

☐$_1$ Yes
☐$_2$ No

3 a. If "Yes", would you describe yourself:

☐1 Cuban
☐2 Mexican
☐3 Puerto Rican
☐4 Other (please specify) ________________

4. How would you describe your primary racial group?

☐1 No Primary Group
☐2 White Caucasian
☐3 Black/African American
☐4 Asian
☐5 American Indian/Alaska Native
☐6 Native Hawaiian/Pacific Islander
☐7 Multi-racial
☐8 Other (please specify) ______________________

5. In which type of housing do you live?

☐1 Residence hall/College dormitory
☐2 House/Apartment/Condominium
☐3 Senior housing (independent)
☐4 Assisted living
☐5 Nursing home
☐6 Relative's home
☐7 Other (please specify) ________________

6. Which category best describes your yearly household income. Do not give the dollar amount, just check the category:

☐1 Less than $5,000
☐2 $5,000 - $9,999
☐3 $10,000 - $14,999
☐4 $15,000 - $19,999
☐5 $20,000 - $29,999
☐6 $30,000 - $39,999

☐7 $40,000 - $49,999
☐8 $50,000 - $59,999
☐9 $60,000 - $69,999
☐10 $70,000 - $99,999
☐11 $100,000 or more
☐12 Do not know for certain
☐13 Do not wish to answer

7. Is English your primary language?

☐1 Yes
☐2 No

7 a. If "No", What is your primary language?

____

8. What is your primary mode of transportation? (Check <u>one</u>)

☐1 Drive my own vehicle
☐2 A friend or family member takes me to places I need to go
☐3 Transportation service provided by where I live
☐4 Use public transportation (e.g., bus, taxi, subway, van services)

<u>Occupational Status</u>

9. What is your primary occupational status? (Check <u>one</u>)

☐1 Work full-time
☐2 Work part-time
☐3 Student
☐4 Homemaker
☐5 Retired
☐6 Volunteer worker
☐7 Seeking employment, laid off, etc.
☐8 Other (please specify)

____

10. Do you currently work for pay?

- ☐1 Yes, Full-time
- ☐2 Yes, Part-time
- ☐3 No

10 a. If "Yes", what is your primary occupation?

____

<u>If retired:</u>

11. What was your primary occupation? ___________________________

12. What year did you retire? ______________________

Health Information

1. In general, would you say your health is:

☐1	☐2	☐3	☐4	☐5
Poor	Fair	Good	Very good	Excellent

2. Compared to other people your own age, would you say your health is:

☐1	☐2	☐3	☐4	☐5
Poor	Fair	Good	Very good	Excellent

3. How satisfied are you with your present health?

☐1	☐2	☐3	☐4	☐5
Not at all satisfied	Not very satisfied	Neither satisfied nor dissatisfied	Somewhat satisfied	Extremely satisfied

4. How often do health problems stand in the way of your doing the things you want to do?

☐1	☐2	☐3	☐4	☐5
Never	Seldom	Sometimes	Often	Always

5. The following items are about activities you might do during a typical day. Does your health now limit you in these activities? Check <u>one</u> box for each type of activity.

	Limited a lot$_1$	**Limited a little**$_2$	**Not limited at all**$_3$
a. Bathing or dressing yourself			
b. Bending, kneeling, or stooping			
c. Climbing **one** flight of stairs			
d. Climbing **several** flights of stairs			
e. Lifting or carrying groceries			

f. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf			
g. Vigorous activities, such as running, lifting heavy objects, or participating in strenuous sports (e.g., swimming laps)			
h. Walking **more than a mile**			
i. Walking **one block**			
j. Walking **several blocks**			

6. Are you on post-menopausal estrogen replacement therapy?

☐1 Yes ☐2 No ☐3 Not applicable

7. For each of the following conditions please indicate if you have ever had that condition in your life, have the condition now at this time or never had the condition. Check <u>one</u> box for each condition.

Condition	**In your lifetime$_1$**	**Now$_2$**	**Never$_3$**
a. Arthritis			
b. Asthma or Bronchitis			
c. Cancer (other than skin cancer)			
d. Diabetes			
e. Epilepsy			
f. Heart Disease			
g. Hearing Impairment			
h. Hypertension			
i. Stroke			
j. Vision Impairment			
k. Other significant illnesses (please list)			

Medication Usage Details

Please list the medical products that you are currently taking. Include medicinal herbs, vitamins, aspirin, etc., as well as prescription medications (copy names from label if possible).

Below is an example of how to fill out the form. If you take Ibuprofen for Arthritis two times a day, you would fill the form out as shown in the example below. There is space for up to eight different medications. If you take more than eight medications regularly, please list the rest on the back of the last page.

Name of Medication	Reason for taking medication	How often do you take this medication? (Please select one)
Example: *Ibuprofen*	*Arthritis*	☒ Daily 2____ times/day ☐ Weekly ______ times/week ☐ Monthly ______ times/month ☐ As Needed

Please turn the page to list your medications

(Note: On the next pages of the questionnaire, tables were provided in the format shown above to list up to eight medications.)

APPENDIX C

ROBOT OPINIONS QUESTIONNAIRE

Imagine that you have the opportunity to use or operate a robot. Please place an X in the response box that best represents your general opinion (we understand that there may be exceptions).

1. My interaction with a robot would be clear and understandable.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

2. I would find a robot useful in my daily life.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

3. Using a robot would enhance my effectiveness in my daily life.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

4. Using a robot in my daily life would increase my productivity.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

5. Using a robot would make my daily life easier.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

6. Using a robot would improve my daily life.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

7. Using a robot in my daily life would enable me to accomplish tasks more quickly.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

8. I would find a robot easy to use.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

9. I would find a robot to be flexible for me to interact with.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

10. It would be easy for me to become skillful at using a robot.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

11. I would find it easy to get a robot to do what I want it to do.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

12. Learning to operate a robot would be easy for me.

☐1	☐2	☐3	☐4	☐5	☐6	☐7
Extremely Unlikely	Quite Unlikely	Slightly Unlikely	Neither	Slightly Likely	Quite Likely	Extremely Likely

APPENDIX D

ROBOT FACIAL APPEARANCE QUESTIONNAIRE

INSTRUCTIONS: Imagine that you were going to be given a robot. You are able to choose what it looks like. The robot would stay with you in your home. It would assist you in performing all the tasks in the home for which you require assistance.

Please answer the following questions about the robot's facial appearance you would want. For each question, please select **one**. There are no right or wrong answers.

1) I would want my robot to have a face.

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

2) I would want my robot to have eyes.

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

3) I would want my robot to have a nose.

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

4) I would want my robot to have a mouth.

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

5) I would want my robot to have lips.

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

6) I would want my robot to have ears.

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

7) **I would want my robot to have hair on its head.**

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

8) **I would want my robot to have soft skin.**

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

9) **I would want my robot to have a round face.**

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

10) **I would want my robot to look exactly like a human.**

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

11) **I would want my robot's face to have female features.**

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

12) **I would want my robot's face to have male features.**

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

13) I would want my robot's face to be unique (i.e., not look like any other robot).

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

14) I would want my robot's face to be attractive.

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

15) I would want my robot's face to be expressive.

❑	❑	❑	❑	❑
Completely disagree$_1$	Slightly disagree$_2$	Neither agree nor disagree$_3$	Slightly agree$_4$	Completely agree$_5$

APPENDIX E

ROBOT FAMILIARITY AND USE QUESTIONNAIRE

For the following robots, please indicate your familiarity in terms of hearing about them, using them, or operating them. Please circle only one option.

Robots	Not sure what this is$_0$	Never heard about, seen, or used this robot$_1$	Have only heard about or seen this robot$_2$	Have used or operated this robot only occasionally$_3$	Have used or operated this robot frequently$_4$
1. Autonomous Car	0	1	2	3	4
2. Domestic/Home robot (e.g., Roomba)	0	1	2	3	4
3. Entertainment/toy robot (e.g., Aibo, Furby)	0	1	2	3	4
4. Manufacturing robot (e.g., robotic arm in factory)	0	1	2	3	4
5. Military Robot (e.g., search and rescue)	0	1	2	3	4
6. Personal Robot 2 (PR2)	0	1	2	3	4
7. Remote presence robot (e.g., Texai, Anybot)	0	1	2	3	4
8. Research robot (e.g., at university or company)	0	1	2	3	4
9. Robot lawn mower	0	1	2	3	4
10. Robot security guard	0	1	2	3	4
11. Space exploration robot (e.g., Mars Rover)	0	1	2	3	4
12. Surgical robot (e.g., da Vinci Surgical System)	0	1	2	3	4
13. Unmanned Aerial Vehicle (UAV)/Drone	0	1	2	3	4

APPENDIX F

ASSISTANCE PREFERENCE CHECKLIST

We are interested in learning about young and older adults' preferences for assistance in performing daily living tasks. In particular, we are looking for opinions about human assistance and robot assistance. When completing this questionnaire, please imagine you need assistance in everyday life with various tasks.

For each of the following tasks, please provide your opinion about your:

- ▪Preference for human assistance
- ▪No preference
- ▪Preference for robot assistance

Assume that the robot could perform the task to the level of a human. Please circle the most appropriate response for your general preference (we understand that there may be exceptions).

On the last page, there is space for you to provide additional comments about your preferences for having robot and human assistance.

If I needed assistance with...	If I needed assistance, I would prefer help from…				
	Only a human$_1$	Prefer a human$_2$	No Preference	Prefer a robot$_4$	Only a robot$_5$
a. Bathing	1	2	3	4	5
b. Being entertained (e.g., playing games, dancing)	1	2	3	4	5
c. Being reminded of appointments	1	2	3	4	5
d. Being reminded of daily activities	1	2	3	4	5
e. Being reminded to take medicine	1	2	3	4	5
f. Brushing teeth	1	2	3	4	5
g. Calling doctors/911	1	2	3	4	5
h. Calling family/friends	1	2	3	4	5
i. Changing light bulbs	1	2	3	4	5
j. Cleaning bathrooms	1	2	3	4	5
k. Cleaning kitchen	1	2	3	4	5
l. Cleaning windows	1	2	3	4	5
m. Controlling for pests/rodents	1	2	3	4	5
n. Deciding what medication to take	1	2	3	4	5
o. Delivering medication	1	2	3	4	5
p. Doing laundry	1	2	3	4	5
q. Eating/feeding myself	1	2	3	4	5
r. Entertaining guests	1	2	3	4	5
s. Exercising	1	2	3	4	5
t. Fetching objects from floor (e.g., remote control) or other room (e.g., drink from refrigerator)	1	2	3	4	5
u. Finding/delivering items (e.g., car keys, glasses)	1	2	3	4	5
v. Gardening/pruning	1	2	3	4	5
w. Getting dressed	1	2	3	4	5
x. Getting information on hobbies/topics of interest	1	2	3	4	5

If I needed assistance with...	If I needed assistance, I would prefer help from…				
	Only a human₁	Prefer a human₂	No Preference	Prefer a robot₄	Only a robot₅
y. Getting information on weather/news	1	2	3	4	5
z. Grocery shopping	1	2	3	4	5
aa. Holding items for you	1	2	3	4	5
bb. Keeping refrigerator clean/stocked	1	2	3	4	5
cc. Learning how to use new technologies	1	2	3	4	5
dd. Learning new physical skills (e.g., dancing)	1	2	3	4	5
ee. Learning new knowledge (e.g., second language)	1	2	3	4	5
ff. Loading/unloading dishwasher	1	2	3	4	5
gg. Maintaining lawn/raking leaves	1	2	3	4	5
hh. Making bed/changing sheets	1	2	3	4	5
ii. Monitoring home/warning about dangers (e.g., fire)	1	2	3	4	5
jj. Monitoring health (e.g., pulse, temperature, blood pressure)	1	2	3	4	5
kk. Opening and closing doors/ drawers	1	2	3	4	5
ll. Painting (e.g., interior/exterior of home)	1	2	3	4	5
mm. Picking up/moving heavy objects (e.g., furniture)	1	2	3	4	5
nn. Preparing meals/cooking	1	2	3	4	5
oo. Reaching for objects	1	2	3	4	5
pp. Reading (e.g., bills, newspaper)	1	2	3	4	5
qq. Rehabilitation exercises	1	2	3	4	5
rr. Repairing plumbing (e.g., fixing leaking faucets)	1	2	3	4	5
ss. Researching medications and health conditions	1	2	3	4	5

If I needed assistance with…	If I needed assistance, I would prefer help from…				
	Only a human[1]	Prefer a human[2]	No Preference	Prefer a robot[4]	Only a robot[5]
tt. Setting the table	1	2	3	4	5
uu. Shaving	1	2	3	4	5
vv. Shopping	1	2	3	4	5
ww. Sorting mail, shredding, throwing away junk mail	1	2	3	4	5
xx. Sweeping/scrubbing/mopping	1	2	3	4	5
yy. Taking medicine	1	2	3	4	5
zz. Taking out trash/recyclables	1	2	3	4	5
aaa. Toileting	1	2	3	4	5
bbb. Turning on/off controls (e.g., switches)	1	2	3	4	5
ccc. Walking	1	2	3	4	5
ddd. Washing dishes by hand	1	2	3	4	5
eee. Washing/combing hair	1	2	3	4	5
fff. Watering plants	1	2	3	4	5

2. If the robot could perform only 5 of the tasks listed on the previous pages, which 5 would you want it to do? (you may list from 0-5 tasks)

1) ______________________________

2) ______________________________

3) ______________________________

4) ______________________________

5) ______________________________

3. Please write any comments about how you answered these questions here:

____

____

____

____

____

____

4. Are there any additional tasks with which you would like robotic assistance? (you may list from 0-5 additional tasks)

1) ______________________________

2) ______________________________

3) ______________________________

4) ______________________________

5) ______________________________

APPENDIX G

PROCEDURE FLOW-DIAGRAM

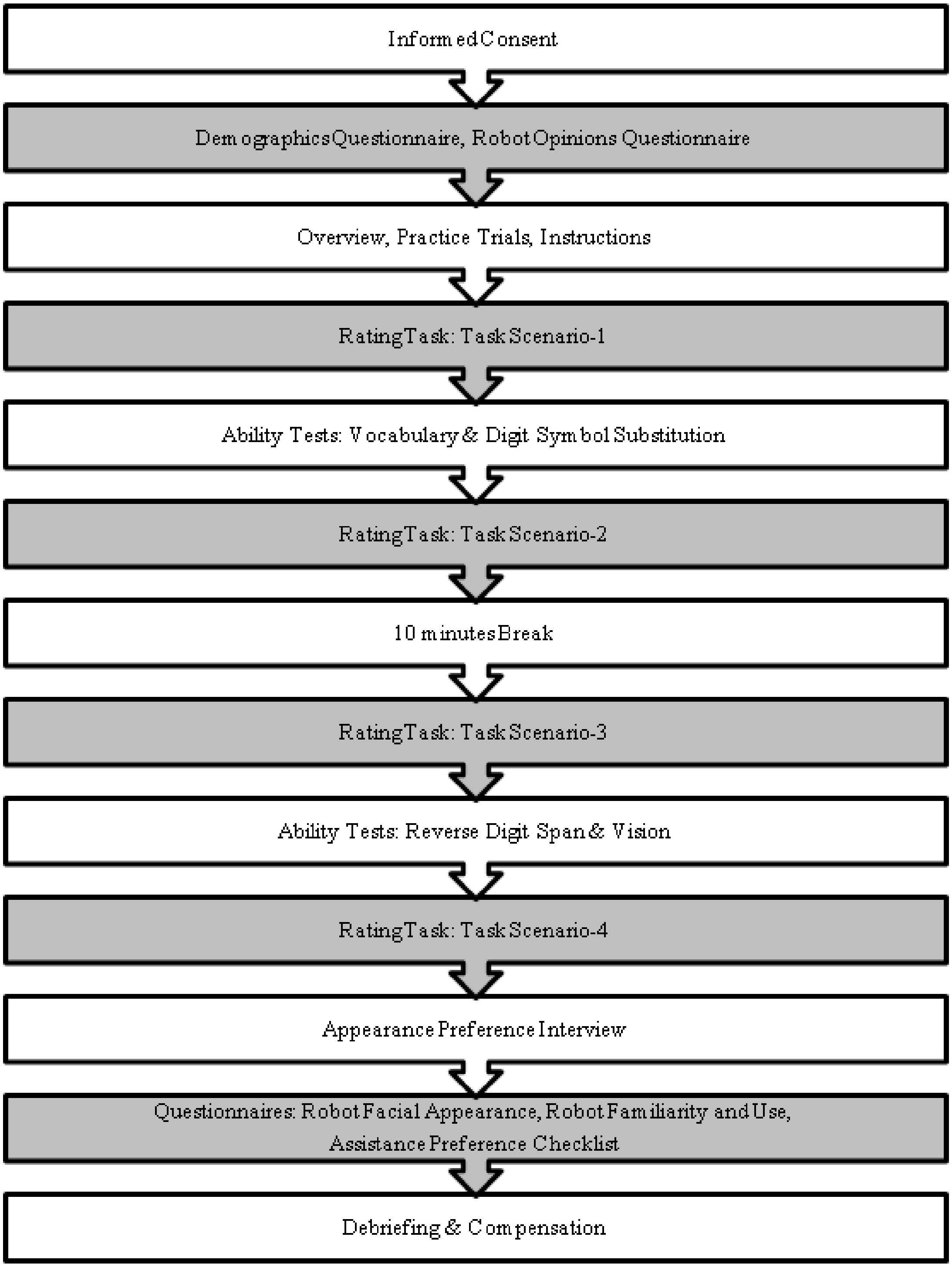


Figure G.1. Procedural flow for the study.

APPENDIX H

COUNTERBALANCING SEQUENCE FOR THE RATING TASK

Table H.1

Counterbalancing Sequence for the Rating Task - Partial Latin Square Design

Sequence	Order of Tasks				Order of Dependent Variables			
1	Social	Chores	Care	Decision	L	A	T	U
2					A	L	U	T
3					T	U	L	A
4					U	T	A	L
5	Chores	Social	Care	Decision	L	A	T	U
6					A	L	U	T
7					T	U	L	A
8					U	T	A	L
9	Care	Decision	Social	Chores	L	A	T	U
10					A	L	U	T
11					T	U	L	A
12					U	T	A	L
13	Decision	Care	Chores	Social	L	A	T	U
14					A	L	U	T
15					T	U	L	A
16					U	T	A	L

L = likeability, A = Anxiety, T = Trust, U = Perceived Usefulness.

The participant in sequence-1 first imagined interaction with a robot in the social task context, followed by chores, personal care, and decision making. For each task, the participant first provided ratings of likeability, followed by anxiety, trust, and perceived usefulness. For each dependent variable, 12 pictures (4 from each of the three levels of humanness) were presented in a random order.

Because there were 32 participants in the two age-groups, 2 participants in each group received one of the 16 counterbalancing balances.

APPENDIX I

INTERVIEW SCRIPT

Imagine that you were selecting a robot that would stay with you in your home. It can assist you in all the tasks you imagined earlier i.e., it can bathe you, perform your daily chores, help you in making investment decisions, and also provide you social companionship.

Now think about the robot's face…

- Which of these two faces would you prefer your robot to have? (show a & b)
- Which of these two faces would you prefer your robot to have? (show c & d)
- Which of these two faces would you prefer your robot to have? (show _ & _)
    - Why would you prefer this face over the other?
        - What is it that you like about this face *(the one selected)?*
        - What is it that you don't like about this face *(the one rejected)?*

Repeat the same pair-wise selection task for

- *Mixed faces (3 comparisons)*
- *Robot faces (3 comparisons)*

Present the most preferred human, mixed and robot face pictures to the participant.

- Which one of these three faces would you prefer your robot to have?
- Why would you prefer this face over the other two?
    - What is it that you like about this face *(the one selected)?*
    - What is it that you don't like about this face *(one of the two rejected)*?
    - What is it that you don't like about this face *(the other one rejected)?*
- Now think specifically about a personal care task, e.g., bathing.
    - Do you have a preference among these faces if the robot helped you in a personal care task?
    - Why would you prefer this face?

- Next, think specifically about a menial task, for example, your daily chores.
  - Do you have a preference among these faces if the robot helped you in a menial task?
  - Why would you prefer this face?
- Now think specifically about a social task, for example, chatting with someone, playing a game with someone, or learning a new skill from someone.
  - Do you have a preference among these faces if the robot helped you in a social task?
  - Why would you prefer this face?
- Finally, think specifically about a decision-making task, for example, deciding where to invest your money.
  - Do you have a preference among these faces if the robot helped you in a decision-making task?
  - Why would you prefer this face?

APPENDIX J

ANALYSIS OF ANXIETY DATA

Described below are reasons for excluding the anxiety data from the main analysis (i.e., from MANOVA). Those data are analyzed separately and presented in this appendix.

1. The term “anxious” was not clearly understood by all participants. Many participants clarified its meaning during the practice-session and could understand it better in terms of “not comfortable”.
2. Although anxiety is considered to be a measure of negative attitudes in the HRI literature (e.g., Heerink et al., 2010; Nomura et al., 2006b), at least two older adults in the present study seemed consider it a positive term. They equated “feeling anxious” to “being eager” about something.
3. Another issue arose due to the participants not paying attention to the scale and/or mistakenly marking their responses in a direction opposite to what they intended (e.g., selecting a "5" (= very much) instead of a "1" (= Not at all)). Nine out of 64 participants reported making some error when marking their anxiety responses. If participants were able to precisely report the mistake they made, their responses were later corrected manually. However, not all participants were unable to trace their errors by memory.

   It is possible that more participants could have made the same mistake without realizing it. At least two older adult participants’ data show that they might have been confused because unlike most participants’ data, their anxiety responses are higher for the pictures they rated high on the other three variables (perceived usefulness, trust and likeability), and lower for the pictures they rated low on the positive variables.

The direction of the anxiety scale was determined so that it is consistent with the direction of other dependent variable scales assessed in the study. (i.e., all scales go from not at all to very much). Reversing the direction for anxiety for could have been another source of confusion for participants. During the early phase of data collection, some modifications were made to overcome the problems noted due to the direction of the anxiety scale. At the end of the practice, participants were reminded again to pay attention to the anxiety scale. Specifically they were told to keep in mind that 1 on the scale means not at all anxious and 5 means very much anxious. However, this did not completely eliminate the problems associated with using anxiety as a measure of perceptions.

Correlation of Anxiety with other DVs

As expected, participants' mean anxiety ratings were negatively correlated with their perceived usefulness, trust, and likeability ratings. These correlations were less strong than those observed among the other three dependent variables. The correlation values are presented in Table J.1.

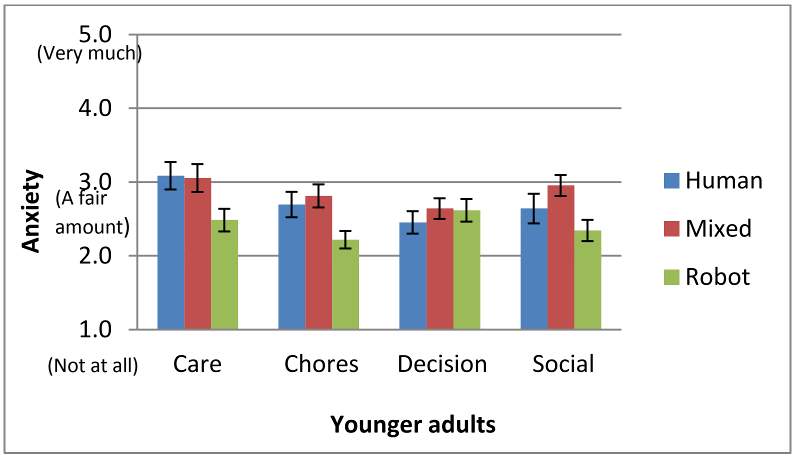
Table J.1

Correlation of Anxiety with Perceived Usefulness (PU), Trust, and Likeability

Age	Variable	r	p
Younger Adults	PU	- .34	0.03
	Trust	- .52	<0.01
	Likeability	- .33	0.06
Older Adults	PU	- .56	<0.01
	Trust	- .55	<0.01
	Likeability	- .46	0.01

How do Perceptions of Anxiety vary as a Function of Age, Humanness and Task?

As observed across the four tasks, and the three levels of humanness (Figure J.1), younger and older adults' mean anxiety ratings fell between 2 (a little) and 3 (a fair amount). Older adults' data were marked by bigger standard error bars indicating more variability in their responses than younger adults. At least part of this variability could have resulted from older adults making more errors when selecting their anxiety responses.



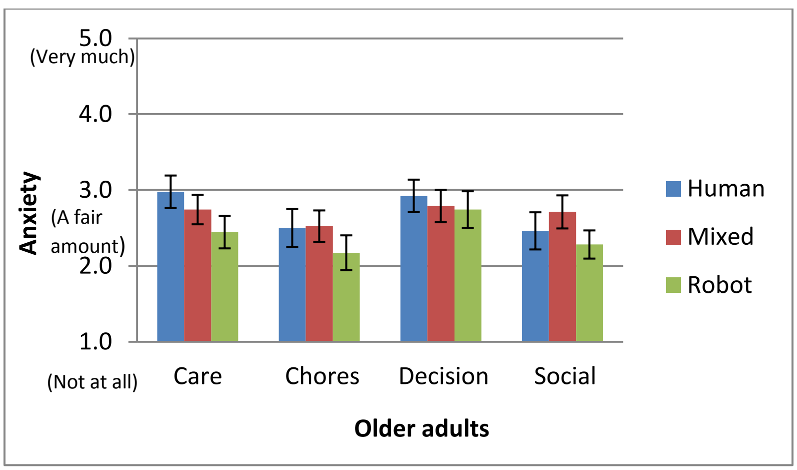


Figure J.1. Mean anxiety ratings by age, humanness, and task. Error bars represent standard error of the mean.

Of the three levels of humanness, least anxiety seemed to be reported for robot appearance. Across the four tasks, participants' reported less anxiety for assistance with chores particularly when the appearance was robot-like. Older adults reported more anxiety for assistance with personal care and decision making, relative to other tasks. Younger adults reported anxiety seemed lower than younger adults for decision making.

An age X humanness X task ANOVA resulted in significant interaction effects of age X task and humanness X task, and a main effect of task. Statistical results of the test are presented in Table J.2.

Table J.2

Summary of the Age X Humanness X Task ANOVA Test Conducted on Anxiety

Effect	F-statistic	DF	p-value	partial η^2
Age X Humanness X Task	0.77	4.58, 284.25	0.56	0.01
Age X Humanness	0.22	1.39, 86.16	0.72	<0.01
Age X Task	3.03*	3, 186	0.03	0.05
Humanness X Task	4.93*	4.58, 284.25	<0.01	0.07
Age	0.15	1, 62	0.70	<0.01
Humanness	3.53	1.39, 86.16	0.05	0.05
Task	5.41*	3, 186	<0.01	0.08

Note: DF is degrees of freedom

A significant interaction effect of age X task on anxiety implied that younger and older adults perceived different levels of anxiety across the four tasks (Figure J.2). Paired-tests were performed to investigate this effect further. The results indicated that both younger and older adults were more anxious in being assisted by a robot with personal care as than with chores (*p's* = 0.01). However, younger adults were more also more anxious to take robot assistance with personal care in comparison to decision making ($p = 0.02$). In case of older adults, assistance with decision making aroused more anxiety in comparison to social task ($p = 0.01$). In summary, older adults reported most anxiety for assistance with decision-making whereas young adults reported most anxiety for assistance with personal care.

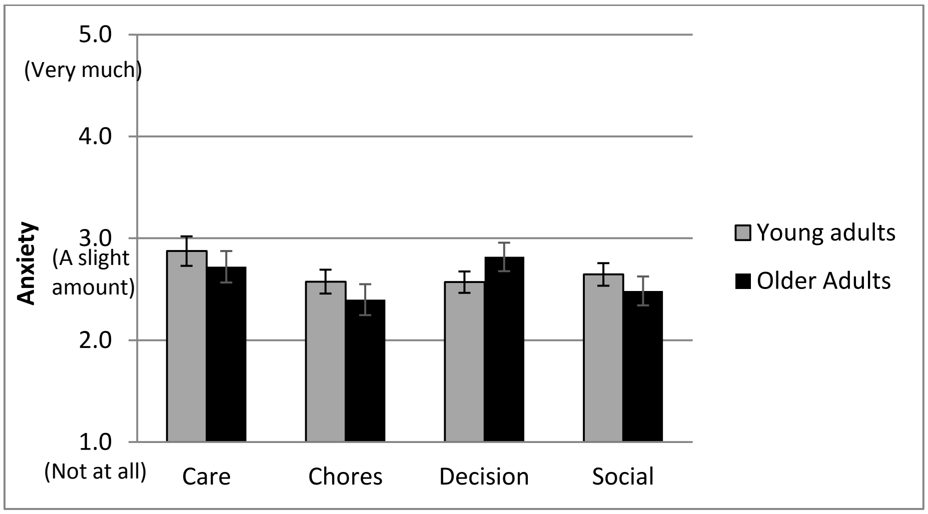


Figure J.2. Mean anxiety ratings by task. Error bars represent standard error of the mean.

A significant humanness X task interaction implied that participants' anxiety across different levels of humanness depended on the task (Figure J.1). Paired t-tests were performed to understand this interaction further. For all tasks except decision making, the mixed appearance evoked more anxiety than the robot appearance (p's < 0.001). For personal care task, even human appearance evoked more anxiety than robot appearance ($p < 0.01$).

Paired t-tests were also conducted to compare the effect of task at every level of humanness. At all three levels of humanness, personal care evoked more anxiety than chores. For human appearance, personal care evoked most anxiety compared to all other tasks. However, for the robot appearance, anxiety was higher for decision-making task in comparison to chores and social task. Thus, overall, comparing across tasks, robot appearance evoked more anxiety when assistance with decision making was imagined. Human appearance evoked most anxiety for assistance with personal care than with other tasks.

The main effect of task was also investigated further by conducting paired t-tests on marginal means. Considering all appearance together, assistance with personal care triggered more anxiety relative to chores ($p < 0.001$) and social task ($p = 0.01$). Moreover, assistance with decision making also evoked more anxiety in comparison to assistance with chores.

Summary of the Effects of Age, Humanness and Task on Anxiety

1. Robot task evoked different levels of anxiety in younger and older adults. Older adults were more anxious for assistance with decision making whereas younger adults seemed more anxious for assistance with personal care.
2. Robot task also moderated the effect of humanness. Human appearance evoked more anxiety for assistance with personal care than with other tasks. Robot appearance evoked more anxiety for assistance with decision making than with chores and social task.
3. Robot task had a main effect on anxiety such that assistance with chores evoked less anxiety than assistance with personal care and decision making. Moreover, assistance with social care also evoked less anxiety in comparison to personal care.

APPENDIX K

MEAN RATINGS FOR ALL FACES

Table K.1

Younger and Older Adults' Mean Ratings for All the Twelve Facial Appearances

Appearance		PU				Trust				Likeability			
		YA		OA		YA		OA		YA		OA	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Human Appearance	**Female1**	2.92	0.88	2.88	1.19	2.59	0.81	2.77	1.19	2.40	0.80	2.65	1.17
	Female2	3.61	0.72	3.41	1.17	3.38	0.95	3.45	1.15	3.16	0.90	3.23	1.18
	Male1	3.07	0.87	2.83	1.20	2.61	0.89	2.69	1.26	2.30	0.83	2.68	1.31
	Male2	2.92	0.77	2.79	1.18	2.72	0.70	2.66	1.22	2.34	0.75	2.56	1.25
Mixed Appearance	**Nexi+Female1**	3.24	0.78	2.62	1.01	2.91	0.89	2.60	1.12	2.70	0.89	2.34	0.94
	Pearl+Female2	2.64	0.85	2.77	1.02	2.54	0.81	2.77	1.03	2.24	0.77	2.70	0.87
	Nao+Male1	2.38	0.90	2.27	1.03	2.20	0.82	2.27	1.03	1.95	0.72	2.02	0.81
	Kobian+Male2	2.68	0.64	2.55	1.07	2.43	0.72	2.42	1.05	2.25	0.68	2.29	0.92
Robot Appearance	**Nexi**	3.30	0.91	3.02	1.25	3.17	0.90	3.00	1.22	3.13	0.98	2.92	1.18
	Pearl	3.18	0.86	2.82	1.09	3.05	0.89	2.88	1.19	3.09	0.91	2.71	1.12
	Nao	3.48	0.92	2.60	1.19	3.63	0.92	2.71	1.23	3.77	0.73	2.63	1.13
	Kobian	2.45	1.01	2.53	1.11	2.39	0.89	2.54	1.16	2.18	0.99	2.26	0.99

*YA = Younger adults; OA = Older adults

APPENDIX L

MEAN RATINGS ON ROBOT FAMILIARITY AND USE QUESTIONNAIRE

Table L.1

Robot Familiarity and Use Questionnaire – Mean Scores and Standard Deviations

Robots	Younger Adults		Older Adults	
	M	SD	M	SD
Autonomous Car	1.84	0.92	1.13	0.83
Domestic/home robot	2.22	0.87	1.78	0.75
Entertainment/toy robot	3.03	0.82	1.78	0.91
Manufacturing robot	2.19	0.40	1.88	0.55
Military robot	1.94	0.25	1.81	0.54
Personal Robot 2 (PR2)	1.00	0.67	1.31	0.82
Remote presence robot	0.84	0.63	0.97	0.78
Research robot	1.69	0.64	1.41	0.71
Robot lawn mower	1.53	0.62	1.56	0.67
Robot security guard	1.28	0.63	1.19	0.65
Space exploration robot	2.00	0.00	1.81	0.47
Surgical robot	1.84	0.37	1.78	0.66
Unmanned Aerial Vehicle	1.63	0.66	1.69	0.69

REFERENCES

Bagozzi, R. P., Davis, F. D., & Warshaw, P. R. (1992). Development and test of a theory of technological learning and usage. *Human Relations, 45*(7), 659-686.

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion 6*(2), 269-278.

Bartneck, C., Kanda, T., Ishiguro, H., & Hagita, N. (2009). My robotic doppelganger - a critical look at the uncanny valley theory. *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN2009* (pp. 269-276). Toyama, Japan.

Beer, J. M., Smarr, C., Chen, T. C., Prakash, A., Mitzner, T. L., Kemp, C. C., Rogers, W. A., (2012). The domesticated robot: Design guidelines for assisting older adults to age in place. *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI).* Boston, MA.

Berry, D. S., & Landry, J. C. (1997). Facial maturity and daily social interaction. *Journal of Personality and Social Psychology*, *72*(3), 570.

Blow, M., Dautenhahn, K., Appleby, A., Nehaniv, C. L., & Lee, D. (2006). The art of designing robot faces - dimensions for human-robot interaction. *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction* (pp. 331-332). Utah, USA.

Broadbent, E., Tamagawa, R., Kerse, N., Knock, B., Patience, A., & MacDonald, B. (2009). Retirement home staff and residents' preferences for healthcare robots. *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication RO-MAN* (pp. 645–650). Toyama, Japan.

Broadbent, E., Tamagawa, R., Patience, A., Knock, B., Kerse, N., Day, K., & MacDonald, B. A. (2011). Attitudes towards health-care robots in a retirement village. *Australasian Journal on Ageing*. doi:10.1111/j.1741-6612.2011.00551.x

Chen, H., Russel, R., & Nakayama, K., & Livingstone, M. (2010). Crossing the ‘uncanny valley’: adaptation to cartoon faces can influence perception of human faces. *Perception, 39*, 378-386.

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit, J. (2006). Factors predicting the use of technology: Findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and Aging*, *21*(2), 333-352.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319-340.

DiSalvo, C. F., Gemperle, F., Forlizzi, J., & Kiesler, S. (2002). All robots are not created equal: the design and perception of humanoid robot heads. *Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (pp. 321-326). London, England.

Eberhardt J. L., Davies P. G., Purdie-Vaughns, V. J., Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science, 17*, 382–386.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological review*, *114*(4), 864.

Ezer, N. (2008). *Is a robot an appliance, teammate, or friend? Age-related differences in expectations of and attitudes towards personal home-based robots.* (Unpublished doctoral dissertation). Georgia Institute of Technology.

Fausset, C. B., Rogers, W. A., & Fisk, A. D. (2009). *Univariate and Multivariate Analysis of Variance: A Primer (HFA-TR-0904).* Atlanta, GA: Georgia Institute of Technology, School of Psychology, Human Factors and Aging Laboratory.

Geller, T. (2008). Overcoming the uncanny Valley. *IEEE Computer Graphics and Applications*, 28(4), 11-17.

Goetz, J., Kiesler, S., & Powers, A. (2003). Matching robot appearance and behavior to tasks to improve human-robot cooperation. *Proceedings of the 2003 IEEE International Workshop on Robot and Human Interaction Communication* (pp. 55-60). Millbrae, CA.

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly, 19*(2), 213-236.

Groom, V., Nass, C., Chen, T., Nielsen, A., Scarborough, J. K., & Robles, E. (2009). Evaluating the effects of behavioral realism in embodied agents. *International Journal of Human-Computer Studies, 67*, 842–849.

Haase, R. F., & Ellis, M. V. (1987). Multivariate analysis of variance. *Journal of Counseling Psychology*, *34*(4), 404-413.

Heerink, M., Kröse, B., Evers, V., & Wielinga, B. (2010). Assessing acceptance of assistive social agent technology by older adults: the Almere model. *International Journal of Social Robotics, 2*(4), 361-375.

Hegel, F., Lohse, M., & Wrede, B. (2009). Effects of visual appearance on the attribution of applications in social robotics. *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 64-71). Toyama, Japan.

Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology*, *89*(3), 416-427.

Hinds, P. J., Roberts, T. L., & Jones, H. (2004). Whose job is it anyway? A study of human–robot interaction in a collaborative task. *Human-Computer Interaction, 19*, 151–181.

Lawton, M. P. (1990). Aging and performance of home tasks. *Human Factors, 32*(5), 527-536.

Lewis, J. D., & Weigert, A. (1985). Trust as a social reality. *Social Forces, 63*, 967-985.

MacDorman, K. F. (2006). Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley. *ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science* (pp. 26–29). Vancouver, Canada.

MacDorman, K. F., Green, R. D., Ho, C.-C., & Koch, C. (2009). Too real for comfort: Uncanny responses to computer generated faces. *Computers in Human Behavior, 25*(3), 695–710.

MacDorman, K. F., & Ishiguro, H. (2006). The uncanny advantage of using androids in social and cognitive science research. *Interaction Studies, 7*(3), 297–337.

Masip, J., Garrido, E., & Herrero, C. (2004). Facial appearance and impressions of credibility: The effects of facial babyishness and age on person perception. *International Journal of Psychology, 39*(4), 276-289.

Mathur, M. B., & Reichling, D. B. (2009). An uncanny game of trust: Social trustworthiness of robots inferred from subtle anthropomorphic facial cues. *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction* (pp.313-314). New York, NY.

McAllister, D. J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal, 38*, 24-59.

Mori, M (1970). Bukimi no tani, *Energy 7*(4), 33–35, translated into English by K.F. MacDorman and T. Minato (2005). Proceedings of the Humanoids-2005 Workshop: Views of the Uncanny Valley, Tsukuba, Japan.

Nomura, T., Kanda, T., Suzuki, T., & Kato, K. (2004). Psychology in human-robot communication: An attempt through investigation of negative attitudes and anxiety toward robots. *Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication* (pp. 35-40).

Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006a). Measurement of negative attitudes toward robots. *Interaction Studies, 7*(3), 437-454.

Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006b). Measurement of anxiety toward robots. *The 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN06)* (pp. 372-377). Hatfield, UK.

Prakash, A., Beer, J. M., Deyle, T., Smarr, C. A., Chen, T. L., Mitzner, T. L., Kemp, C. C. & Rogers, W. A. (2013, March). Older adults' medication management in the home: how can robots help? *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction* (pp. 283-290). Tokyo, Japan.

Riggio, R. E., Widaman, K. F., Tucker, J. S., & Salinas, C. (1991). Beauty is more than skin deep: Components of attractiveness. *Basic and Applied Social Psychology, 12*, 423-439.

Rogers, W. A., Meyer, B., Walker, N., & Fisk, A. D. (1998). Functional limitations to daily living tasks in the aged: A focus group analysis. *Human Factors, 40*(1), 111-125.

Schneider, E., Wang, Y., & Yang, S. (2007). Exploring the uncanny valley with Japanese video game characters. *Proceedings of Situated Play, DiGRA 2007 Conference* (pp. 546-549).

Shipley, W. C. (1986). Shipley institute of living scale. Los Angeles: Western Pyschological Services.

Smarr, C., Fausset, C. B., & Rogers, W. A. (2011). *Understanding the Potential for Robot Assistance for Older Adults in the Home Environment (HFA-TR-1102).* Atlanta, GA: Georgia Institute of Technology, School of Psychology, Human Factors and Aging Laboratory.

Smarr, C. A., Prakash, A., Beer, J. M., Mitzner, T. L., Kemp, C. C., & Rogers, W. A. (2012). Older adults' preferences for and acceptance of robot assistance for everyday living tasks. *Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting* (pp. 153-157). Santa Monica, CA: Human Factors and Ergonomics Society.

Snellen, H. (1868). Test-types for the determination of the acuteness of vision (4th ed.). London: Williams & Norgate.

Thompson, J. C., Trafton, J. G., & McKnight, P. (2011). The perception of humanness from the movements of synthetic agents. *Perception, 40*(6), 695-704.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308*(5728), 1623-1626. doi: 10.1126/science.1110589

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science, 46*(2), 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425-478.

Walster, E., Aronson, V., Abrahams, D., Rottman, L. (1966). Importance of physical attractiveness in dating behavior. *Journal of Personality and Social Psychology, 4*(5), 508-516.

Walters, M. L. (2008). *The design space for robot appearance and behaviour for social robot companions.* (Unpublished doctoral dissertation). School of Computer Science, University of Hertfordshire.

Walters, M. L., Syrdal, D. S., Dautenhahn K., Boekhorst, R., Koay, K. L. (2008). Avoiding the uncanny valley – robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. *Journal of Autonomous Robots, 24*(2), 159-178.

Wechsler, D. (1997). Wechsler adult intelligence scale III. (3rd Ed.) San Antonio. TX: The Psychological Corporation.

Weiss, H. M., & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. In B. M. Staw & L. L. Cummings (Eds.), *Research in organization behavior* (Vol. 19, pp. 1-74). Greenwich, CT: JAI Press.

Wu, Y. H., Fassert, C., & Rigaud, A. S. (2012). Designing robots for the elderly: Appearance issue and beyond. *Archives of Gerontology and Geriatrics*, *54*(1), 121-126.

Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?* Boulder, CO: Westview Press.

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass*, *2*(3), 1497-1517.

Zebrowitz, L. A., & Rhodes G. (2004). Sensitivity to “bad genes” and the anomalous face overgeneralization effect: Cue validity, cue utilization, and accuracy in judging intelligence and health. *Journal of Nonverbal Behavior, 28*, 167–185.